Paidós Básica

Últimos títulos publicados

52. S. Toulmin y J. Goodfield - *El descubrimiento del tiempo*
53. P. Bourdieu - *La ontología política de Martin Heidegger*
54. R. Rorty - *Contingencia, ironía y solidaridad*
55. M. Cruz - *Filosofía en la historia*
56. M. Blanchot - *El espacio literario*
57. T. Todorov - *Crítica de la crítica*
58. H. White - *El contenido de la forma*
59. F. Rella - *El silencio y las palabras*
60. T. Todorov - *Las morales de la historia*
61. R. Koselleck - *Futuro pasado*
62. A. Gehlen - *Antropología física*
63. R. Rorty - *Objetividad, relativismo y verdad*
64. R. Rorty - *Ensayos sobre Heidegger y otros pensadores contemporáneos*
65. D. Gilmore - *Hacerse hombre*
66. C. Geertz - *Conocimiento local*
67. A. Schütz - *La construcción significativa del mundo social*
68. G.E. Lenski - *Poder y privilegio*
69. M. Hammersley y P. Atkinson - *Etnografía. Métodos de investigación*
70. C. Solís - *Razones e intereses*
71. H.T. Engelhardt - *Los fundamentos de la bioética*
72. E. Rabossi y otros - *Filosofía de la mente y ciencia cognitiva*
73. J. Derrida - *Dar (el) tiempo 1. La moneda falsa*
74. R. Nozick - *La naturaleza de la racionalidad*
75. B. Morris - *Introducción al estudio antropológico de la religión*
76. D. Dennett - *La conciencia explicada. Una teoría interdisciplinar*
77. J.L. Nancy - *La experiencia de la libertad*
78. C. Geertz - *Tras los hechos*
79. R.R. Aramayo, J. Muguerza y A. Valdecantos - *El individuo y la historia*
80. M. Augé - *El sentido de los otros*
81. C. Taylor - *Argumentos filosóficos*
82. T. Luckmann - *Teoría de la acción social*
83. H. Jonas - *Técnica, medicina y ética*
84. K.J. Gergen - *Realidades y relaciones*
85. J.R. Searle - *La construcción de la realidad social*
86. M. Cruz (comp.) - *Tiempo de subjetividad*
87. C. Taylor - *Fuentes del yo*
88. T. Nagel - *Igualdad y parcialidad*
89. U. Beck - *La sociedad del riesgo*
90. O. Nudler (comp.) - *La racionalidad: su poder y sus límites*
91. K.R. Popper - *El mito del marco común*
92. M. Leenhardt - *Do kamo*
94. T. Eagleton - *Ideología*
95. M. Platts - *Realidades morales*

Mark Platts

Realidades morales

Ensayo sobre psicología filosófica

PAIDÓS
México • Buenos Aires • Barcelona

Instituto de Investigaciones Filosóficas
Universidad Nacional Autónoma de México

Título original: *Moral Realities. An Essay in Philosophical Psychology*
Publicado en inglés por Routledge, 1991

Traducción de Ana Isabel Stellino y Antonio Zirión
Cubierta de Mario Eskenazi

1a edición, 1998

Quedan rigurosamente prohibidas, sin la autorización escrita de los titulares del «Copyright», bajo las sanciones establecidas en las leyes, la reproducción total o parcial de de esta obra por cualquier medio o procedimiento, comprendidos la reprografía y el tratamiento informático, y la distribución de ejemplares de ella mediante alquiler o préstamo público.

D.R. © 1991 Mark de Bretton Platts
D.R. © 1991 by Routledge
D.R. © de la presente edición en castellano:
Universidad Nacional Autónoma de México

Coeditan: Instituto de Investigaciones Filosóficas, UNAM;
Circuito Mario de la Cueva, Ciudad de la Investigación en Humanidades,
Ciudad Universitaria, 04510, México, D.F.,
Tels.: 622-7241, 622-7438. Fax: 665-4991

Ediciones Paidós Ibérica, S.A.,
Mariano Cubí, 92 - 08021 Barcelona, y

Editorial Paidós Mexicana, S.A.,
Rubén Darío 118, 03510, col. Moderna, México, D.F.,
Tels.: 579-5922, 579-5113. Fax: 590-4361

ISBN 968-853-874-2

Derechos reservados conforme a la ley
Impreso en México - Printed in Mexico

A la memoria de Juan Manuel Romo Santos

SUMARIO

PREFACIO 11
INTRODUCCIÓN 13

PRIMERA PARTE

1. Falsas concepciones del deseo 25
   1.1 Una concepción russelliana del deseo 25
   1.2 Sobre el deseo como poder activo 27
   1.3 El deseo y las teorías causales de la acción 34
   1.4 Deseo e intención 36
   1.5 Explicación de la acción 39
   1.6 Caracterizaciones de deseabilidad 40
   1.7 «Deseos animales» 42

2. Las distinciones del deseo 47
   2.1 Necesidades, carencias, deseos 47
   2.2 Caracterizaciones de deseabilidad y percepciones de deseabilidad 52
   2.3 ¿Otra concepción falsa? 54
   2.4 Objetos del deseo profundamente oscuros 58
   2.5 El deseo mismo 62
   2.6 Algunas distinciones iniciales 63
   2.7 Algunas cuestiones iniciales 67
   2.8 Refinamiento de las distinciones iniciales 68
   2.9 Deseo y deseabilidad 73
   2.10 Deseo y placer 74
   2.11 El simple querer 76

3. Valores 85
   3.1 Un escepticismo cauteloso 85
   3.2 Del querer al valorar 87
   3.3 El internalismo del valor 93
   3.4 Del valorar al querer: deseos potenciales no motivados 96

3.5 Del valorar al querer: deseos potenciales motivados 100
3.6 El sentido de la subjetividad, de la dignidad, del mérito ... 107
3.7 ¿Se trasciende lo subjetivo sensible? 114

SEGUNDA PARTE

4. Hecho y acción en la teoría moral de Hume 125
4.1 El argumento maestro 125
4.2 La doctrina del aislamiento 133
4.3 Algunas dificultades 144
4.4 Causación razonable de deseos 146
4.5 Moralidad y acción 151
4.6 Hecho y sentimiento 157

5. El alcance de la moralidad 163
5.1 Las cuestiones internas 163
5.2 Discrepancias morales 173
5.3 Conceptos y concepciones 175
5.4 Cuestiones externas 183
5.5 Un relativismo contemporáneo 188
5.6 Otro relativismo contemporáneo 196
5.7 ¿Perspectivas para un relativismo razonable? 199

6. Críticos de la moralidad 209
6.1 *Terra incognita* 209
6.2 Pompa, vanidad e impaciencia 213
6.3 ¿Un caballo que no baila? 216
6.4 ¿Todas las abejas han perdido su aguijón? 222
6.5 El amplio manto de la sociabilidad 224
6.6 Una hoja de parra moral 227
6.7 La congruencia inglesa 230
6.8 Si Dios ha muerto 235
6.9 El topo y el enano 241

Posfacio 249
Bibliografía 251
Índice analítico 257

PREFACIO

Thomas Love Peacock le atribuía a un viejo amigo suyo la opinión de que publicar un libro sin prefacio es como entrar a un salón sin hacer una reverencia. Michael Dummett piensa que encontrarse con un libro sin prefacio es como llegar a cenar a una casa y ser conducido directamente al comedor. Así que ésta es mi ligera reverencia, junto con la invitación al lector a que se sirva un trago (fuerte, pues la cena puede resultar indigesta).

En la redacción de este ensayo me he basado en varias de mis publicaciones (aunque en todos los casos he desarrollado o modificado las opiniones expuestas en ellas). Por ello agradezco a los editores correspondientes la autorización para incluir materiales de los siguientes artículos:

1. «La moralidad, la personalidad y el sentido de la vida», *Diálogos*, no. 117, 1984.
2. «The Object of Desire», *Crítica*, XVII, no. 50, 1985.
3. «Desire and Action», *Noûs*, XX, no. 2, 1986.
4. «Hume and Morality as a Matter of Fact», *Mind*, XCVII, no. 386, 1988.
5. «¿Tiene algún porvenir la filosofía moral?», *Revista Latinoamericana de Filosofía*, XIV, 1988.
6. «Introducción» y «Hume: La moralidad y la acción», en Mark Platts (comp.), *La ética a través de su historia*, UNAM, México, 1988.
7. «The metaphysics of morals», en R.R. Verma y P.K. Sen (comps.), *The Philosophy of P.F. Strawson*, Nueva Delhi, Indian Council of Philosophical Research, 1994.

También estoy en gran deuda con Martha Sasía por la paciencia y habilidad con que convirtió un borrador ilegible, aunque supuestamente en inglés, en un manuscrito legible, y al Sistema Nacional de Investigadores de México por el apoyo que hizo posible mi estancia en este país. Pero, sobre todo, estoy en deuda con John McDowell y Paul Snowdon por sus comentarios sobre la penúltima versión de este ensayo. Ellos evitaron que incurriera en muchos errores que, de otro modo, habrían quedado im-

presos en estas páginas, y me hicieron muchas sugerencias útiles para mejorar el texto.

Como en mi primer libro –al cual una vez alguien en México se refirió en forma impresa, muy gratamente, como *Waste of Meaning* (Desperdicio de significado)–,* considero las ideas de otros autores sólo en la medida en que ello contribuye a la comprensión de las ideas que aquí presento. Me habría gustado hacerme eco del pensamiento de Collingwood y decir que menciono a otros sólo *honoris causa*, pero podría haberse creído que eso significaba sumarle daño al insulto. Así que he intentado limitar esas referencias al mínimo. (Incluso había pensado agregar una segunda bibliografía de obras que, si bien no se citan en el texto principal, estoy seguro de que han influido en mí; pero el riesgo de pecar de *doble* omisión me hizo abandonar la idea.) Sin duda, el lector reconocerá ciertas influencias que no menciono; sólo espero que nadie se sienta ofendido.

Una clase de omisión enteramente distinta es la de toda consideración de cuestiones morales de «primer orden», de «ética práctica». Puesto que es muy posible que esto resulte decepcionante para algunos, quizás sea preciso señalar en este punto que más bien deberían tenerse por afortunados. Mi concepción del mundo es sombría, y mi opinión sobre la eficacia de la discusión de temas de «primer orden» está lejos de ser optimista; cuando hay seres humanos de por medio, mi tendencia natural consiste en suponer que la luz al final del túnel es un tren que se nos viene encima. No obstante, admito de buen grado que he gozado de un privilegio: llegué a conocer a alguien que, sin tener conciencia de ello, abrió los ojos de sus semejantes a las cosas de valor aparentemente pequeñas que hay en este mundo y, así, hizo de su paso por él algo digno de ser vivido y compartido, no sólo soportado –y menos todavía rechazado con ideas confusas y carentes de significación–. No fue un hombre famoso ni un santo, pero todos los que lo conocieron vieron sus vidas inmensamente enriquecidas. Este libro es para él.

M. de B.P.
Instituto de Investigaciones Filosóficas,
Universidad Nacional Autónoma de México,
México, D.F.

* *Ways of Meaning*, versión en castellano: *Sendas del significado. Introducción a una filosofía del lenguaje*, trad. Cecilia Hidalgo y Eduardo A. Rabossi, Instituto de Investigaciones Filosóficas, UNAM/Fondo de Cultura Económica, México, 1992.

INTRODUCCIÓN

> Señor, si caigo en un río, un hombre sencillo se zambullirá y me sacará; pero un filósofo me observará con la mayor serenidad, me considerará a la luz de la teoría de los proyectiles y, calculando el grado de fuerza con que he golpeado la superficie, la resistencia del líquido, la velocidad de la corriente y la profundidad del agua en ese sitio particular, asegurará de la manera más impecable en qué parte del lodo del fondo será probable que se me encuentre en cualquier intervalo de tiempo dado desde el momento en que me sumergí por primera vez.
>
> THOMAS LOVE PEACOCK

Karl Kraus sostenía que la moralidad era una enfermedad venérea que tenía como fase primaria la virtud, como fase secundaria el aburrimiento y como fase final la sífilis. No se necesita reflexionar mucho para percatarse de que Kraus pensaba en las ortodoxias relativas a la «moralidad sexual» que predominaban en su época y su país, así como en la hipocresía que implicaba su condición misma de ortodoxias. Resulta obvio también que estaba haciendo una crítica ferozmente *moral* del contenido de esas ortodoxias y de esa hipocresía. Tampoco es muy difícil hallar ejemplos de otras críticas de vasto alcance, pero aun así morales, de las ortodoxias morales en culturas o sociedades específicas. En realidad, hasta es relativamente fácil identificar casos de críticas de la moralidad en su conjunto, de *todas* las moralidades, fundadas en ciertos valores no morales que los críticos suscriben. Parece que el tema de la defectuosidad de la moralidad ha sido siempre un compañero íntimo de tales críticas. Y por ello difícilmente resultaría sorprendente por sí mismo tropezar ahora con la idea de que la moralidad está en un estado de grave desorden, que está hecha un lío.

Tampoco deberíamos asombrarnos al descubrir que esa idea va acompañada de otra: la de que la *filosofía* de la moralidad también está hecha un lío. Supongamos que la moralidad es, de una u otra forma, radical e irreparablemente defectuosa. Supongamos además que, habiéndonos

percatado de esto, llegáramos a aceptar que, dentro de lo humanamente posible, tenemos que abandonarla por completo. Supongamos incluso —aunque no es esencial para lo que aquí se quiere decir— que en efecto la abandonáramos. Bien podríamos entonces continuar todavía con el estudio filosófico de la rechazada institución de la moralidad, tal como, por así decirlo, podríamos continuar con el estudio filosófico de la magia o la religión. La continuación de ese estudio filosófico podría parecer, en el mejor caso, un asunto algo endeble. Sin embargo, es importante reconocer que, en tal contexto, el estudio filosófico de la moralidad se asemejaría mucho más al de la religión que al de la magia. En el caso de la religión subsiste, por ejemplo, una tarea que en parte es filosófica y que tiene un innegable interés: la de identificar las necesidades, los deseos y los intereses que, por lo menos supuestamente, la institución religiosa satisface; hacer referencia a ellos podría servir así para explicar en gran medida la persistencia de esa institución. Y quedaría luego la tarea subsecuente de considerar y evaluar las opciones alternativas que se podrían encaminar a satisfacer esas necesidades, deseos e intereses. Exactamente las mismas tareas surgirían después del hipotético abandono total de la institución de la moralidad: en parte corresponderían al objeto de estudio de la entonces futura filosofía moral, en parte al objeto de estudio de sus sucesores.

De suerte que si la moralidad se hallara en efecto en un estado radical e irreparablemente defectuoso, ello acarrearía consecuencias serias y problemáticas para la práctica de la filosofía moral. Pero hay otra presunta analogía entre la moralidad y la religión que posee una importancia mucho mayor para nuestros actuales propósitos. En el caso de las situaciones defectuosas tanto de la religión como de la filosofía de la religión, muchos han creído que existe una conexión en sentido opuesto; es decir, muchos han pensado que el estado defectuoso de la religión se debe precisamente a que los filósofos de la religión no han logrado ofrecer fundamentos filosóficos coherentes de las creencias y prácticas religiosas. Y muchos han afirmado lo mismo en el caso de la moralidad: han tenido la convicción de que la moralidad se encuentra en un lío esencialmente filosófico. Pero yo pienso que esta convicción encierra una gran complejidad, y es lo que intentaré exponer ahora.

Al abordar cualquier institución humana, cualquier práctica humana, surge de inmediato el problema de su *identificación*. Consideremos el caso de la institución de la ciencia. Como caracterización inicial —tosca pero eficaz— podríamos decir que esta institución consiste por lo menos en las siguientes actividades: investigaciones científicas empíricas, como los

experimentos de laboratorio; formulación de explicaciones científicas; postulación de leyes y teorías científicas; publicación de artículos y libros científicos; participación en conferencias científicas; enseñanza de la ciencia; administración de subvenciones para la investigación científica, etc., etc., etc....

Una vez que consideramos una institución de esta clase, podemos imaginar, en términos generales, dos tipos distintos de teorías acerca de la institución. Una teoría es *interna* de la institución, de la práctica, cuando se afirma que quienes participan en ésta hacen lo que hacen porque creen, aunque quizás tácitamente, en la teoría de que se trata. Una teoría es *externa* a la institución si no se satisface esa condición. ¿Pero cuál es exactamente el contenido real de esa condición y por tanto de esa distinción? Una propuesta inicial podría ser que al menos parte de ese contenido se puede formular de la manera siguiente: quien propone una teoría interna sostiene en realidad que, si los participantes de la institución de que se trata llegaran a rechazar esa teoría, entonces *ceteris paribus* suspenderían su práctica, abandonarían la institución.

Hay que señalar dos cuestiones antes de proseguir. En primer lugar, en el ejemplo dado, la caracterización inicial de la institución de la ciencia puede parecer tosca e ineficaz, y esta impresión se debe quizás al hecho de que la expresión «científico» se emplea repetidas veces dentro de esa caracterización. Y en segundo lugar, podría parecer que la frase «*ceteris paribus*» usada en el intento por explicar siquiera una parte del contenido de la distinción entre las teorías internas y externas volvería inútil esa explicación. La idea que se oculta tras el empleo de dicha frase está clara: incluso después de que se llega a rechazar la teoría interna de cierta institución, los participantes podrían permanecer dentro de ella, podrían continuar con la misma práctica, por cualquiera de un número indefinido de razones —falta de imaginación, hábito, seguridad económica duradera, etc., etc., etc....— Es imposible ofrecer *a priori* una lista completa de motivaciones por las cuales los seres humanos pueden ingresar a una actividad determinada o permanecer dentro de ella. Pero, en consecuencia, parece que, pese a todo lo dicho hasta aquí, la distinción entre teorías internas y externas carece de una diferencia empírica.

No obstante, supongamos que esas inquietudes logran disiparse de un modo razonable. Y supongamos también que la teoría interna de la institución de la ciencia es, al menos parcialmente, una teoría filosófica. Entonces, en estas circunstancias, los fracasos de los filósofos para hallar fundamentos filosóficos que sustenten esa teoría podrían poner de manifiesto el estado crítico de la ciencia misma: la ciencia podría hallarse

inmersa en un lío esencialmente filosófico. En estas circunstancias, no es que los filósofos hayan incursionado en un terreno en el que no tenían derecho de estar; más bien sería que están donde de todos modos se los necesita. Pero, en estas circunstancias, son incapaces de satisfacer esa necesidad.

Volvamos ahora a la institución de la moralidad. Ésta se compone por lo menos de lo siguiente: nuestros pensamientos y juicios morales (evaluaciones morales específicas de una variedad de tipos de cosas, evaluaciones morales generales de esas mismas cosas, juicios morales prácticos dirigidos a acciones posibles); nuestras prácticas morales, como castigar y recompensar; nuestras emociones morales, como la gratitud y la culpa. Con respecto a esta institución, podemos contemplar la posibilidad de que tenga una teoría interna que sea, por lo menos en parte, filosófica; y luego podemos contemplar la posibilidad ulterior de que esa teoría resulte filosóficamente indefendible. Cuando contemplamos estas posibilidades, nos encontramos considerando una situación en la cual el trabajo de los filósofos morales podría poner de manifiesto el lío en que la moralidad está inmersa. En tales circunstancias, la moralidad necesita una fundamentación filosófica, pero *ex hypothesi* la moralidad en esas circunstancias no tiene lo que necesita.

Una de estas circunstancias es la siguiente: que la teoría interna de la moralidad resulte filosóficamente indefendible. Ahora bien, si esto es así podría serlo tan sólo debido a las insuficiencias de los filósofos morales del momento; puede suceder que estos filósofos sean incapaces de defender la teoría —o que tengan una concepción errónea de lo que sería una defensa exitosa de la teoría—. De tal suerte, no tiene que ser un problema *filosófico* el hecho de que existan problemas morales arduos, de que el discurso moral se utilice a menudo con fines no morales, de que no haya unanimidad en cuestiones morales o de que abunde la inmoralidad a nuestro alrededor; más bien, la idea general que aquí se necesita es la de que hay algún objetivo o alguna condición que la moralidad debe satisfacer para que resulte defendible y que no satisface. Pero entonces tiene que considerarse la posibilidad de que el problema resida en el objetivo o la condición que los filósofos han impuesto a una defensa exitosa de la moralidad. Así pues, bien podría ser que los mismos filósofos morales se encontraran en un estado de grave desorden. Pero supongamos que no es así, y ahora, finalmente, supongamos que los filósofos morales del momento han descubierto una prueba filosófica convincente de que la teoría interna de la institución de la moralidad es en verdad indefendible —que es falsa o incoherente o sin sentido o carente de fundamento—.

Entonces, en estas circunstancias, el escepticismo *filosófico* respecto de la moralidad sería apropiado. En estas circunstancias, la moralidad se halla –y es razonable creerlo así– en una confusión esencialmente filosófica, confusión que los filósofos son capaces de apreciar con precisión. Pero las circunstancias en cuestión son en verdad complejas.

Tales escepticismos filosóficos respecto de la moralidad han tomado muchas formas, pero hay dos que son especialmente instructivas. Una deriva de la larga, liosa historia que ha desembocado en el pensamiento y la práctica morales del presente. Este pensamiento y esta práctica son resultado de diferentes influencias o factores históricos cuya compatibilidad no es obvia en absoluto: desde las nociones originadas en la Grecia clásica y las ideas distintivas del cristianismo, pasando por las preocupaciones de la Ilustración, hasta por lo menos las contribuciones del liberalismo, el utilitarismo, el socialismo y el marxismo del siglo XIX (véase MacIntyre 1981: 6–11). Una reflexión sobre esa historia parecería sugerir fácilmente que nuestro pensamiento y nuestra práctica morales actuales no son más que una mezcolanza de muy dudosa racionalidad. La segunda clase de escepticismo filosófico respecto de la moralidad se ocupa de varios elementos aparentemente distintivos del pensamiento y la práctica morales, en los que encuentra, a lo sumo, una coherencia muy dudosa: la idea del libre albedrío, por ejemplo, o el pensamiento de que las razones morales para actuar son razones que los agentes poseen al margen de sus deseos y creencias –son «razones externas» para actuar– (véase Williams 1981a; 1985: cap. 10).

Desde luego, hay importantes diferencias entre las formas resultantes del escepticismo filosófico respecto de la moralidad. Es probable, por ejemplo, que ofrezcan respuestas enteramente distintas a las siguientes preguntas: ¿siempre ha estado la moralidad en un lío o las supuestas deficiencias de su estado son de origen más reciente? Y ¿existe alguna alternativa a abandonar por completo la institución de la moralidad?, ¿puede salvarse algo de las ruinas presentes? No obstante, estas formas del escepticismo filosófico respecto de la moralidad tienen también algo en común: todas suponen que resulta (relativamente) fácil identificar la teoría interna de nuestra institución de la moralidad. Pero quisiera insistir en que esa identificación está, en realidad, muy lejos de ser fácil.

Un ejemplo de la historia reciente de la filosofía moral servirá para indicar la dificultad del caso. Hace algunos años Philippa Foot intentó, con muy buena fortuna, poner en tela de juicio la coherencia de la idea de un «imperativo categórico» como el que emplean los filósofos de la tradición kantiana (Foot 1972a); y más recientemente, Bernard Williams

ha emprendido una tarea similar en relación con lo que llama la idea de una «razón externa» para actuar (Williams 1981a). Desde luego, había diferencias sutiles entre los objetivos y los métodos de ambos filósofos, a pesar de lo cual, la semejanza general, por lo menos en las ideas que atacaban, resultaba clara. Lo interesante para nuestros propósitos presentes estriba en que, en términos generales, mientras la señora Foot parece haber supuesto que estaba criticando una concepción errónea de los filósofos morales sobre la naturaleza de la moralidad, Williams parece haber supuesto que estaba criticando un error interno del pensamiento moral mismo. La primera atacaba una mera tesis de los filósofos; el otro, un elemento de la institución misma de la moralidad. La cuestión de cuál de estos filósofos tenía la apreciación más verídica del carácter de su objetivo común, revierte en el problema de cuál es la teoría de la que verdaderamente puede afirmarse que es interna de nuestro pensamiento y nuestra práctica morales. El carácter elusivo de ese problema queda atestiguado por la diferencia, todavía no resuelta, entre los filósofos citados.

La dificultad general así indicada podría guardar una relación asombrosamente estrecha con el problema, muy distinto en apariencia, de la identificación de la institución de la moralidad. Una idea razonablemente atractiva consiste en que lo único que habría que tratar de hacer para llegar a resolver esa cuestión es contrastar elementos *específicos* de nuestro pensamiento y nuestra práctica morales con otras cosas *específicas*, es decir, enfocarse en contrastes doblemente específicos. Pero también podría considerarse otra posibilidad, a saber: sostener que la identidad de la institución de la moralidad está determinada, en los términos más generales, por el asunto de la teoría de la que pueda afirmarse verdaderamente que es interna de ella. Y entonces se podría introducir además la tesis de que la preocupación que instó a incluir la problemática cláusula *ceteris paribus* era espuria; pues aun cuando los agentes parezcan continuar comportándose como participantes de la institución después de haber rechazado su teoría interna, ese rechazo muestra que no pueden seguir siendo como antes participantes de la *misma* institución. Si se presenta ahora la objeción de que no se ha considerado el problema de, por así decirlo, la identificación «externa» de la institución, la respuesta apropiada bien podría ser afirmar que no hay razones para suponer, en general, que exista un problema semejante que deba ser considerado. Desde luego, los términos con que se caracterice a la teoría interna de la institución de la moralidad tienen que ser de tal índole que se presten a una aplicación empírica; pero no hay razón para creer que esa condición de la aplicabilidad empírica deba interpretarse desde el punto de vista de alguna concepción

conductista de las manifestaciones externas de esa teoría, ni siquiera solamente en términos que excluyan el uso de las nociones morales mismas en las descripciones de las posibles manifestaciones externas. No es preciso que intervenga aquí ninguna interpretación reduccionista de la idea de una teoría interna, de suerte que si la exigencia de que se considere el problema de la identificación externa de la institución es la exigencia de semejante interpretación, tenemos todo el derecho de rechazar esa exigencia.

Si en verdad puede decirse que hay una teoría por lo menos parcialmente filosófica interna de la institución moral, esa teoría sería el objeto de una *metafísica descriptiva de la moralidad*, en un sentido de «metafísica descriptiva» similar al que le da Strawson (véase Strawson 1959; 1985). La construcción de semejante metafísica sería un intento por describir las estructuras y los rasgos más generales supuestamente incluidos en nuestro pensamiento y nuestra práctica morales: un intento por poner al desnudo las conexiones y prioridades conceptuales más generales contenidas en ese pensamiento y esa práctica. La construcción de semejante metafísica de la moralidad representaría, en cierto sentido, una tarea más modesta que la construcción de alguna otra metafísica descriptiva, ya que una vez que se aprecie la cualidad distintiva de la institución de la moralidad, parece improbable que la metafísica resultante arroje mucha luz sobre «el contraste entre lo que es *inevitable* en la estructura del pensamiento humano y lo que es contingente y modificable» (Hampshire 1959: 9; las cursivas son mías). Aun así, ese toque de modestia puede parecer engañoso en lo que concierne al carácter general de la empresa de construir una metafísica descriptiva de la moralidad: pese a todo lo que hasta aquí se ha dicho, parece que todavía no tenemos una idea muy clara, si acaso tenemos alguna, de cómo acometer esa empresa. De hecho, algunas consideraciones adicionales sirven para aumentar la incertidumbre respecto de este punto.

De nada nos servirá en esta empresa un examen de lo que la gente dice *acerca* de la moralidad. Esto no sólo se debe a que cualquier teoría interna de la moralidad podría ser aceptada por los agentes morales únicamente de manera tácita, ni sólo a que es probable que las opiniones expresadas representen meramente opiniones morales de «primer orden»; también se debe a que son demasiados los factores extraños que pueden entrar en juego al determinar lo que dice la gente sobre la moralidad —cortesía, provocación, autoengaño, ideas filosófico-culturales apresuradas, etc., etc., etc.— Más preocupante, quizás, es el hecho de que está muy lejos de resultar obvio que el metafísico descriptivo de la moralidad pueda obtener

mucha ayuda del examen del *uso* que la gente hace del vocabulario moral. J.L. Mackie aparentemente pensaba que es más o menos evidente que nuestro pensamiento y nuestro discurso morales pretenden ser objetivos (Mackie 1977: cap. 1); pero, en oposición a esto, Simon Blackburn ha intentado mostrar que cualquier rasgo de nuestro uso del lenguaje moral que se imagine que es la característica definitoria de dicha postura supuestamente objetiva, puede ser reproducido en su uso del lenguaje moral por un subjetivista consumado, *grosso modo* del tipo de Hume (Blackburn 1984: cap. 6). Así, por ejemplo, ni el discurso sobre creencias, verdades y hechos morales, ni el compromiso con cierto principio de ambivalencia sirven para distinguir al objetivista del subjetivista. Desde luego, cualquier conclusión general que se siga de la línea de razonamiento de Blackburn es vulnerable a una posibilidad pasada por alto. No obstante, la moraleja tentativa que quiero extraer a partir del estado actual de este debate consiste en que es probable que un examen de hechos pertenecientes a nuestro uso del vocabulario moral no desempeñe sino un papel relativamente indirecto en la empresa de identificar la teoría interna de nuestra institución de la moralidad. Pero entonces, ¿cómo debemos acometer esa empresa?

En la prolongada y tortuosa historia de la metafísica de la moral, puede detectarse a veces una respuesta característica a esa pregunta. En los términos más esquemáticos, se trata de la propuesta de que el mejor enfoque para identificar la teoría de la que verdaderamente pueda decirse que es interna de nuestro pensamiento y nuestra práctica morales consiste en la investigación filosófica de nuestra psicología moral. Desde el punto de vista de la estrategia, ello equivale a decir que la filosofía moral se entiende mejor como parte de la psicología filosófica. Pero es esencial distinguir esa recomendación estratégica de los resultados a los que presuntamente llegan ciertos filósofos cuando la adoptan. El más importante de estos filósofos, Hume, pensaba que el despliegue de esa estrategia nos conduciría a resultados subjetivistas, de acuerdo con los cuales el material metafísico y aparentemente objetivo de la moralidad se *reduce* a meras realidades psicológicas. Sin embargo, es posible aprobar la recomendación estratégica y dudar al mismo tiempo de esos resultados, pues cabe sospechar que, según los términos de la recomendación, una descripción verídica de las realidades psicológicas correspondientes apoyará, en términos descriptivos, una metafísica de la moralidad diametralmente opuesta a la de Hume (por lo tanto, cabe sospechar que las realidades psicológicas descritas como tales carecen de la independencia conceptual, respecto del

pertinente material metafísico objetivo de la moralidad, que se requiere para la tarea *reduccionista*).

Este ensayo constituye una exploración parcial de esa posibilidad –la estrategia de Hume sin los resultados de Hume–, una exploración que espero que, por lo menos, sirva para dar algo de cuerpo al esqueleto de esa posibilidad que acabo de describir. Cuando, pensando en esta posibilidad, comencé a tomar la filosofía moral en serio, me topé con un gran escollo: el concepto de deseo. Este concepto ocupa un lugar central no sólo en la discusión filosófica sobre la moralidad, sino también en la discusión filosófica de una gama de temas notablemente amplia, al grado de que cualquier esperanza de efectuar un «análisis» riguroso del concepto parece destinada a fracasar. Sin embargo, ello no excluye la posibilidad ni reduce la necesidad de una especulación menos formal encaminada a sacar a la luz algo que por lo menos se aproxime a una concepción convincente de lo que es el deseo. Malinterpretar un concepto tan central no puede menos que causar estragos en nuestra comprensión de lo que pensamos acerca del puesto de la mente en el mundo.

La primera parte de este ensayo es, principalmente, el resultado de mis intentos por obtener alguna claridad respecto de ese concepto. El primer capítulo se apoya en uno de los casos raros en que un gran filósofo ha ofrecido una declaración explícita sobre su concepción del deseo. Procedo así en este capítulo con el fin de eliminar algunos malentendidos casi naturales y muy difundidos acerca de la naturaleza del deseo. El segundo capítulo, más constructivo, es consecuencia del intento de presentar una concepción más convincente sobre lo que es el deseo y sobre el puesto que ocupa en la vida mental y activa de los seres humanos; tal intento se realiza desde una perspectiva determinada por algunas distinciones de suma importancia entre las clases de deseo. Por último, en el tercer capítulo, se exponen los resultados de esa discusión más constructiva con el propósito de presentar lo que podría llamarse, en términos un poco arcaicos y grandilocuentes, una *Teoría General del Valor*: una metafísica descriptiva, general y sistemática del valor. Para ese punto ya deberá haberse aclarado, en líneas generales, la manera como quiero divorciar la estrategia humeana de los resultados subjetivistas humeanos.

En la segunda parte del ensayo, esa metafísica descriptiva del valor se afina un poco más mientras intento desarrollarla como una parte de la empresa de construir una metafísica descriptiva parcial de la moralidad. Esta segunda parte del ensayo es, en general, mucho menos constructiva que la primera. Al tratar de arrojar luz sobre la teoría interna de la institución de la moralidad, abordo esa institución desde varios ángulos

diferentes: desde una consideración de las concepciones de dicha institución que han sostenido los más grandes de los filósofos que se han opuesto a las concepciones que yo mismo quiero defender (capítulo 4); desde el examen de uno de los problemas, el referente al relativismo moral, que supuestamente surgen en relación con esa institución y que se han tratado mucho en obras filosóficas recientes (capítulo 5); y desde una reflexión sobre las concepciones de algunos filósofos que han criticado esa institución, o que por lo menos se considera que la han criticado (capítulo 6). El resultado consiste, a lo sumo, en una metafísica descriptiva *parcial* de la moralidad, y esto en más de un sentido: sólo presto atención de manera constante y directa a algunas características del pensamiento y la práctica morales; poco o nada se dice acerca de los complementos que aún se necesitarían si se quisiera alcanzar la exhaustividad; y poco o nada se dice explícitamente sobre lo que se requeriría para que con esos complementos se lograra un buen acabado, hablando en términos descriptivos. Si el resultado es también parcial en el sentido peyorativo de esa expresión, es una cuestión que queda desde luego a juicio del lector.

Hay una acusación, sin embargo, que me gustaría rechazar desde el principio. No considero que mi interés por los asuntos *descriptivos* sea la manifestación de una preferencia conservadora y meramente temperamental, ni tampoco la manifestación de una serena pasión por el cálculo de las maneras en que la moralidad nos tiene atascados en el fango. Acaso la teoría de la que verdaderamente puede afirmarse que es interna del pensamiento y la práctica morales es defectuosa en uno u otro sentido; acaso esa teoría es defectuosa en un sentido característicamente filosófico; acaso deberíamos reevaluar ahora todos nuestros conceptos y prácticas morales y cuestionar la institución entera en la cual esos conceptos y esas prácticas encuentran su lugar. Acaso, acaso, acaso... Pero esas pretensiones críticas y esos proyectos revisionistas o revolucionarios adquirirán contenidos diáfanos y motivaciones transparentes solamente cuando la teoría de la que en verdad pueda afirmarse que es interna del pensamiento y la práctica morales haya sido realmente identificada.

PRIMERA PARTE

No hay prejuicio más natural para el hombre que el de concebir que la mente tiene alguna similitud con el cuerpo en sus operaciones.

THOMAS REID

Aunque en un amplio círculo de gente cada una de las personas pueda estar sentada en las rodillas de quien está detrás de ella, ésta no es una hazaña que dos o tres personas solas puedan lograr.

GARETH EVANS

CAPÍTULO 1

FALSAS CONCEPCIONES DEL DESEO

> «Entre le désir et l'action, monsieur, il y a place pour le respect». La phrase était drôle, bien que peu claire.
>
> GUY DE MAUPASSANT

> Única regla básica para los expertos en mujeres: limítese sólo a explicar por qué hizo ello lo que ya hizo, pues eso le ahorrará el problema de explicar por qué no hizo lo que usted dijo que haría.
>
> REX STOUT

1.1 UNA CONCEPCIÓN RUSSELLIANA DEL DESEO

En *The Analysis of Mind* (1921),* Russell presentó una exposición muy característica del deseo: concisa, inteligente, honesta y casi perversamente imaginativa. Entre otras muchas aseveraciones –algunas de las cuales tal vez le interesaban más–, sostenía las siguientes:

(1) La concepción del deseo que tiene el sentido común está radicalmente equivocada.
(2) La concepción del deseo que tiene el sentido común lo considera, en parte, un sentimiento específico hacia cierta imagen.
(3) La concepción del deseo que tiene el sentido común lo considera, en parte, una atracción del futuro más que un impulso desde el presente –un jalón y no un empujón–.
(4) El estudio de los animales no humanos constituye, en muchos sentidos, la mejor preparación para el análisis del deseo.
(5) El principal motor en la acción conforme al deseo es una sensación de incomodidad.
(6) El deseo es una ley causal de nuestras acciones: un impulso o una tendencia a la acción, un poder que influye en las acciones.

Pueden encontrarse, creo, ideas parecidas en las nociones tácitas de muchos filósofos acerca de qué es el deseo; representa una virtud por parte

* Versión en castellano: *Análisis del espíritu*, trad. Eduardo Prieto, Buenos Aires, Paidós, 1962.

de Russell el haber hecho explícita su concepción. No obstante, adaptando una frase suya, el deseo constituye un asunto sobre el cual, si no estoy equivocado, sólo se pueden alcanzar concepciones verdaderas por medio de una inversión casi completa de la opinión *filosófica* irreflexiva corriente (y de la opinión reflexiva de Russell).

Pero antes, una advertencia. En este caso, mi discusión no pretende guardar una correspondencia simple con los pormenores del uso de la expresión específica «desear» y sus cognadas, ni siquiera con los pormenores del uso del más humilde «querer» y sus cognadas. Algunos aspectos de la relación entre mi análisis y los hechos en torno a ese uso se aclararán en el transcurso de la discusión, pero como indicación inicial, el tema de esta última está anclado por el siguiente pensamiento: siempre que un agente ϕa intencionalmente, desea (o quiere) ϕar.

Así entendidas (o mal entendidas), algunas de las afirmaciones de Russell pueden descartarse de inmediato. Quizás la víctima más obvia es la número 5, la afirmación universal de que el principal motor de la acción conforme al deseo es una sensación de incomodidad. Pues en muchísimos casos ordinarios de acción intencional, está claro que no está presente ninguna sensación de incomodidad que mueva al agente a actuar: los casos normales de cruzar la calle, abrir el periódico, cerrar la puerta, hablar. Y a menos que el sentido común esté notablemente ciego ante ese hecho, se revela un error similar en la número 2, la afirmación de que la concepción del deseo del sentido común lo considera, en parte, un sentimiento específico hacia cierta imagen. ¿Por qué el sentido común se negaría para sí mismo o bien la variedad fenomenológica que se manifiesta en los casos de deseo, o bien el conocido vacío fenomenológico que ocurre en muchos casos ordinarios de acción intencional conforme al deseo? ¿Por qué no puede el sentido común reconocer tanto las distinciones entre los casos de deseo como el hecho evidente, por ejemplo, de que la incidencia del deseo *sentido* depende en buena parte de la medida en que la acción conforme al deseo se vea obstruida por dificultades psicológicas o físicas?

Si continuamos considerando con un mínimo de benevolencia al sentido común, veremos que la afirmación 2 encierra otro error, ya que en muchos casos ordinarios de acción intencional también es obvio que el agente no necesita tener ninguna imagen de la acción futura o de cierta situación resultante futura. En algunos casos de planeación y deliberación, las imágenes del futuro pueden desempeñar algún papel, y la incidencia de tales imágenes puede ser mayor cuando se presentan obstrucciones y dificultades. Pero la afirmación universal atribuida al sentido común

en el punto número 2 es tan claramente falsa que resulta difícil entender cómo podría el sentido común incurrir en semejante error.

El error que Russell atribuye al sentido común en la afirmación 3 –la concepción del deseo como un jalón desde el futuro más que como un empujón desde el presente– es de una naturaleza completamente distinta. Pues aquí no está involucrada una mera falsificación fenomenológica de las experiencias del deseo y de la acción conforme al deseo, sino más bien una curiosa metafísica de la causación: una metafísica en la cual lo que ahora todavía no existe causa algo. Pero la atribución de Russell de semejante metafísica al sentido común se ve socavada por una exposición mucho más verosímil de la concepción del sentido común acerca de este asunto. Con base en esta exposición, la concepción del asunto consiste en que el «principal motor» en la acción conforme al deseo es algo que existe *ahora* –a saber, el deseo mismo–, junto con la idea de que una especificación completa de esa cosa implica la especificación de su *contenido proposicional* característico. Ese contenido será, por lo menos generalmente, una descripción de una posible acción o estado de cosas futuro; pero su especificación no *requerirá* que se atribuya a quien desea una imagen de esa acción o estado de cosas futuro. Mi deseo es que cruce yo la calle, y así lo hago; pero la causa de que así lo haga no es algo que todavía no existe, ni resulta necesario que yo tenga ahora una imagen de ese algo, sino que más bien consiste meramente en el deseo que ahora tengo con el contenido proposicional especificado. Sin duda, hay oscuridades filosóficas en esa concepción del asunto que deben ser aclaradas; y quizás eso, *la* concepción del deseo del sentido común, esté equivocada. Pero de ser así, no se debe ni a una falsificación obvia de la fenomenología del deseo ni a una oscura metafísica de la causación.

Por fortuna, en la discusión russelliana del deseo hay errores más instructivos, algunos de los cuales guardan relación con los errores más superficiales que hasta aquí se han mencionado.

1.2 Sobre el deseo como poder activo

Russell sostenía (como en la afirmación 6) que el deseo es una ley causal de nuestras acciones, un impulso o una tendencia a la acción, un poder que influye en las acciones. Al sostener esto, Russell se colocaba dentro de una tradición filosófica casi universal, en la cual, como es de esperarse, se pueden detectar sutiles diferencias de opinión y de énfasis. Quiero ahora examinar algunos de los principales miembros de la familia de las concepciones sobre el deseo que componen esa tradición. Algunos miembros

presentan insuficiencias obvias, pero todos, según creo, tienen defectos. Si puede hacerse valer esta afirmación de manera sistemática, se establecerá el resultado de que no es sino una falsa concepción pensar que el deseo consiste esencialmente en un poder activo de la mente, es decir, en términos generales, pensar que consiste esencialmente en una disposición o tendencia a actuar para tratar de producir el estado de cosas deseado.

Los deseos se atribuyen a los agentes como parte de la tarea de hacer que sus actos nos resulten inteligibles. Más específicamente, las atribuciones de deseo son componentes clave de las racionalizaciones de las acciones intencionales. Por ejemplo, en lo que podría parecer el caso más sencillo, la acción de que se trata se considera, por medio de la racionalización, productora «directa» de la situación deseada. Pero, desde luego, existen innumerables formas más complejas en que puede considerarse que un deseo es algo conforme a lo cual se actúa.

Quien se concentre exclusivamente en esa gama de usos dados al concepto de deseo puede verse conducido de modo natural a abrazar la concepción del deseo a la que dio aguda expresión el creador de Don Quijote: «todas las veces que el deseo de alguna cosa se enciende en nuestros corazones, luego nos mueve a seguirla y a buscarla, y buscándola y siguiéndola, a mil desordenados fines nos conduce» (*La Galatea*, Libro IV).

El error lo introduce aquí ese enfoque. Las atribuciones de deseos a los agentes no sólo se hacen como parte de la elaboración de descripciones racionalizadoras de sus acciones intencionales, de suerte que se considere que los agentes actúan conforme a los deseos que se les han atribuido. Aun cuando todas las atribuciones de deseos deban «fundarse» en ciertos aspectos de la conducta del agente, todavía hay muchas formas en las que sus deseos pueden ponerse de manifiesto en su conducta que difieren de su actuar conforme a ellos. Por lo tanto, no todos los deseos exigen que se actúe conforme a ellos. Y esto es así también puesto que el agente en cuestión puede no tener idea alguna de cómo buscar y alcanzar el objeto de su deseo; es decir, puede no tener ninguna creencia respecto de cuáles de los cursos de acción que tiene ante él harían más probable que el estado de cosas deseado tuviera lugar. En verdad, no sería posible actuar conforme a todos los deseos. El conjunto de deseos que el agente tiene puede ser de tal índole, en virtud de sus contenidos, que éste crea correctamente que es imposible actuar conforme a todos ellos de un modo razonable, en el sentido especial de actuar conforme a todos ellos con la esperanza razonable de que se produzcan todos los estados de cosas deseados. Tal

vez son más comunes los conjuntos de deseos de cuyos miembros se cree correctamente que resultan en este sentido incompatibles por alguna otra circunstancia contingente; un ejemplo interesante es el que ofrece la brevedad de la vida en relación con cualquier acervo considerable de deseos. Para la mayoría de las personas es imposible la mayor parte del tiempo que todos sus deseos se ajusten a la concepción cervantina.

Los modos más frecuentes de intentar rectificar esa concepción frente a las evidentes fallas mencionadas recurren, entre otras cosas, tanto a cierta idea sobre las creencias del agente acerca de lo que está dentro de su poder producir, como a una noción sobre la fuerza comparativa de sus deseos. De tal suerte, que un deseo llegue en efecto a originar alguna acción depende de las creencias del agente y de la naturaleza y la fuerza de sus otros deseos. Se puede desarrollar este pensamiento de varias formas distintas; al considerar tales formas, el asunto se simplifica al concentrar la atención más estrechamente en la noción de que un agente o quiera hacer algo o desee hacerlo. Así, por ejemplo, se dejan de lado los problemas —de los que cualquier teoría general del deseo debe forzosamente ocuparse— que surgen del querer una balandra, de Quine, y el deseo de amor, de Prufrock.

Recientemente, Donald Davidson ha defendido el siguiente principio sobre la noción de que un agente quiera hacer algo:

> P1. Si un agente quiere hacer *x* más de lo que quiere hacer *y* y se cree libre de hacer ya sea *x* o *y*, hará intencionalmente *x* si hace *x* o *y* intencionalmente. (Davidson 1969: 23)

Una objeción común a ese principio queda ilustrada por el siguiente tipo de aparente contraejemplo. Quiero visitar a Juan más de lo que quiero visitar a Tomás, y correctamente me creo libre de visitar a uno u otro; pero Tomás vive cerca, mientras que Juan vive lejos en las afueras; así que visito a Tomás, no a Juan, intencionalmente. No hay en ese resultado nada que pueda parecer ni siquiera mínimamente desconcertante; pero la idea general que se esconde tras la objeción es que sí lo hay para Davidson, por la sencilla razón de que su principio P1 se concentra exclusiva y erróneamente, por así decirlo, en los beneficios esperados de la acción contemplada, descuidando de esta manera la cuestión de los costos esperados de la acción en cuanto al tiempo y el esfuerzo requeridos, la incomodidad causada y otros factores.

Podemos rebatir esta objeción tan pronto como se plantee. Cualquier teórico tendrá que distinguir entre querer hacer *x* más que *y* haciendo

abstracción de la consideración de los costos, y querer hacer *x* más que *y* tomando en cuenta los costos esperados de hacer una u otra cosa. Puedo querer visitar a Juan más que a Tomás en los términos del primer caso, haciendo abstracción de los costos, y querer visitar a Tomás más que a Juan una vez que tomo en cuenta los costos esperados. Así, el principio P1 de Davidson queda protegido de inmediato contra la clase de contraejemplo aparente planteada –y queda en verdad interpretado de un modo más natural– cuando se consideran las expectativas del agente respecto de los costos de las acciones contempladas como algo que está reflejado en las fuerzas comparativas de sus deseos, que es interno de ellas. De manera similar, los diferentes grados de confianza que experimenta el agente respecto de los beneficios y costos de las acciones pueden entenderse como internos de la fuerza de sus deseos. Y para aquellos a quienes les desagrade mi importación de jerga pseudoeconómica a la descripción del caso, existen varias formas más o menos naturales de describir éste y otros casos semejantes. El siguiente diálogo (que creo que se debe a Philippa Foot) nos ofrece otro ejemplo:

A: Quiero dejarte ahora.
B: Creí que te simpatizaba. ¿Cómo puedes *querer* dejarme?

Por «querer» la persona A entiende, en términos generales, *querer en vista de todas las circunstancias*; la persona B entiende, aproximadamente, *querer por o en sí mismo*. El anclaje inicial de mi tema –siempre que un agente ϕa intencionalmente, desea (o quiere) ϕar– revela con claridad que aquí lo que me interesa en general es la primera clase de uso. Uno puede hacer algo intencionalmente aun cuando no quiera hacerlo *por o en sí mismo* –quizás sea la menos desagradable de las opciones que se le presentan–.

Lo que parecería requerir una modificación sustancial de la formulación del principio P1 de Davidson, contra una mera clarificación de su interpretación, es el deseo de que, en aras de la amplitud, se aplique a casos en que el grado de confianza del agente en su capacidad para ejecutar las acciones contempladas varía entre una y otra de tales acciones. Esa consideración resulta a menudo crítica en los casos en que está involucrada la dimensión temporal; comprender muchos de esos casos también exige que se tome en cuenta cualquier creencia que tenga el agente, y la fuerza de cualquier creencia semejante, respecto de la posibilidad de que ejecute más de una de las acciones contempladas. Ambas consideraciones, junto con los grados de expectativa del agente en lo que toca a beneficios y costos, son de especial importancia para comprender de manera realista la contemplación del agente, no sólo de acciones aisladas,

sino de planes de acción que se extienden hasta un futuro lejano. Pero para nuestros propósitos actuales, podemos restringir nuestra atención a los casos del tipo más sencillo.

En este momento quiero más comprar un hipopótamo de porcelana mañana por la mañana que comprar en este momento una botella de champaña. Creo que estoy en libertad de hacer una cosa u otra, pero también creo que si hago lo segundo no podré hacer lo primero. Sin embargo, compro en este momento una botella de champaña, y lo hago intencionalmente, frustrando así en apariencia, y creyendo que lo frustro, el deseo más fuerte. ¿Cómo explicar mi acción?

Dado lo que se estipuló previamente sobre la internalización de la expectativa, así como del grado de ésta, de los costos relativos a la fuerza de los deseos del agente, no se puede esperar ahora ayuda alguna por ese lado. La forma más natural de comprender muchos casos de este tipo consiste en considerar que la creencia del agente de que las dos acciones contempladas son incompatibles es una convicción no cabal; estimamos que cree o espera, por irrazonable que sea, que «surgirá algo» que le permitirá actuar, todavía racionalmente, conforme a su deseo más fuerte cuando llegue el momento de tal acción. Otra posibilidad (potencialmente conectada con la primera), que tiene significación también cuando la dimensión temporal no es importante, consiste en sostener que el grado de creencia del agente en su capacidad para ejecutar la acción es apreciablemente diferente para cada acción contemplada. Esa diferencia puede ser resultado de alguna creencia, por parte del agente, relativa al carácter no temporal de las acciones y de sí mismo; o, si no, podría ser específicamente resultado de las supuestas diferencias temporales que reflejan, digamos, las creencias generales del agente relativas a la incertidumbre del futuro (o de su futuro). Otra posibilidad más, también potencialmente conectada, es que el agente aminore un tanto su deseo (en otras circunstancias) más fuerte y dirigido a un futuro más distante, por abrigar ciertas dudas acerca de si todavía tendrá ese deseo o, por lo menos, si lo seguirá teniendo con su fuerza intacta, cuando llegue el momento de actuar.

La última clase de consideración puede tomarse razonablemente como algo interno de la fuerza actual de los deseos del agente. Determinar si las otras posibilidades de explicación mencionadas, que invocan las creencias, admiten ser tratadas en términos de una internalización comparable es una cuestión espinosa. Sin embargo, dejando a un lado lo espinoso, podríamos contemplar la aparente posibilidad de los casos del tipo ejemplificado, en los cuales nos vemos obligados a desechar todas las clases de explicación posible que acabamos de esbozar. Quedaría por lo menos otra

posibilidad de explicación: aun cuando todas las demás consideraciones han sido tomadas en cuenta, el agente está todavía más influido por la dimensión temporal. Esto también podría expresarse de modo natural respecto de la internalización. Aun cuando se hayan tomado en cuenta todas las demás consideraciones, el deseo conforme al cual el agente cree que puede actuar en un momento dado es, en virtud de ese hecho, más fuerte que el deseo conforme al cual el agente sólo cree que puede actuar en un futuro más distante. Esto podría ser verdad únicamente respecto de los deseos particulares que ahora estamos considerando, o podría ser verdad respecto de todos los deseos que el agente tiene en el momento actual, o podría ser verdad respecto de todos los deseos del agente a lo largo del tiempo. Independientemente de las otras consideraciones mencionadas, un agente puede tener cierta propensión, constante o no constante, y de una intensidad estable o no estable, a aminorar sus deseos actuales pero dirigidos a un futuro más distante debido a que están dirigidos a un futuro más distante.

El hecho de que se puedan explicar los casos del tipo ejemplificado no significa que en ellos las acciones del agente estén más allá de todo reproche. Cada una de las posibilidades de explicación mencionadas va acompañada de una posibilidad de crítica de la conducta del agente distinta; y cada una corresponde a un tipo distinto de deliberación protopráctica.

Por importantes que sean esas consideraciones para comprender, y quizás para enmendar, el principio P1 de Davidson, no parece que necesiten una modificación sustancial de la concepción de Davidson sobre el querer hacer algo. Creo que se puede apreciar el aspecto que en esencia es el más discutible cuando preguntamos por qué Davidson agrega a su principio la cláusula «si hace *x* o *y* intencionalmente». Una razón obvia, surgida por el nivel de generalidad en el que Davidson maneja su discusión, es la necesidad de darle acomodo a la posibilidad de que el agente tenga un deseo todavía más fuerte conforme al cual se crea capaz de actuar. Una segunda razón de obviedad comparable es la necesidad de dejar un margen para los casos en que la creencia del agente en su capacidad para hacer lo que más quiere sea falsa. Y una tercera razón, en apariencia distinta para tal agregado, se deriva de los casos en que un agente, aunque correctamente se cree capaz de hacer lo que más quiere, no logra en esa ocasión ejercer correctamente esa capacidad: intenta la acción apropiada pero, esa vez, su ejecución es defectuosa.

Si eso agota las principales consideraciones que determinan el agregado de Davidson de la cláusula en cuestión, entonces me parece que su principio es simplemente poco convincente. En este punto será útil reflexionar sobre lo que pueden denominarse «casos aparentes de un único deseo». Estoy sentado ante mi escritorio, mirando distraídamente por la ventana. El cuarto está lleno de humo, me escuecen los ojos y mi incomodidad aumenta a cada momento que pasa. Por ello, quiero abrir la ventana. Creo que tal acción está dentro de mi alcance actual. Y, no obstante, permanezco sentado, mirando morosamente por la ventana. Tal escena no es poco frecuente, por lo menos en mi vida. Pero, ¿se sigue que no quiero en realidad abrir la ventana? ¿O se evita esa consecuencia solamente sosteniendo que mi creencia de que soy capaz de abrir la ventana es falsa o bien que tengo, contra todas las apariencias, otro deseo más fuerte conforme al cual estoy en ese momento por lo menos tratando de actuar? Sin duda no es así: sencillamente puede suceder que el único deseo pertinente en mi situación actual, y por lo tanto mi deseo pertinente más fuerte, no me mueva a la acción ni me mueva a intentar la acción, y que ni siquiera origine una intención apropiada de actuar. Se podría decir que el deseo no logra involucrar a mi voluntad.

Al considerar la descripción de tales casos, es importante tener en mente una doctrina de viejo arraigo filosófico de la cual el mismo Davidson nos dice que «posee cierto aire de evidencia», a saber: «que en la medida en que una persona actúa intencionalmente, actúa, según lo expresa Santo Tomás, a la luz de un bien imaginado» (Davidson 1969: 22). Ahora bien, en muchos casos de un único deseo aparente podemos dar razón de la inacción del agente haciendo referencia a alguna característica pasajera de su estado de ánimo, de su estado emocional o, en términos más generales, de su vida mental en el momento en que la acción es posible. En otras ocasiones, la razón puede darse aludiendo a algún rasgo más persistente de la personalidad del agente. Diré que estas razones ofrecen, respectivamente, explicaciones mentales y personales de la inacción del agente, sin que con ello se quiera sugerir que existe aquí una distinción tajante entre los tipos de explicación. De las razones que podemos presentar, algunas introducen, por lo menos tácitamente, una referencia a otros deseos del agente; y de estas razones, algunas pueden revelar que su inacción es intencional, que es resultado de la influencia apropiada de otro deseo supuestamente más fuerte. Pero, ¿acaso debe proceder así todo intento aceptable de dar razón de la inacción de un agente en un caso aparente de un único deseo? La dificultad es obvia: en muchos casos, simplemente no existe ninguna especificación convincente de un «bien

imaginado» a la luz del cual el agente permanecería intencionalmente inactivo. Dicho de un modo más sencillo, con frecuencia no existe ninguna especificación convincente de un deseo (más fuerte) conforme al cual el agente está actuando intencionalmente. Incluso dentro del contexto de una creencia correcta del agente sobre sus capacidades adecuadas, el principio P1, según se entiende aquí, incluye una visión excesivamente simplificada de cuándo su deseo más fuerte desembocará en una intención de actuar correspondiente, ya no digamos en la acción misma.

1.3 El deseo y las teorías causales de la acción

Hasta aquí el argumento puede reconciliarse con una concepción del deseo que lo considere esencialmente un poder activo de la mente, con la afirmación, por ejemplo, de que atribuirle a una persona el querer algo equivale a decir que está dispuesta a tratar de conseguirlo. Pues todo lo que se nos ha revelado es la existencia de algunas ambigüedades y complejidades en cualquier especificación convincente de las circunstancias en las cuales se ejerce ese poder o se realiza esa disposición. Pero la siguiente afirmación continúa incólume:

> P1* Si un agente se cree capaz de hacer lo que más quiere, entonces normalmente intentará hacerlo.

La afirmación no carece de contenido, pues la fuerza del *normalmente* estriba en expresar un compromiso con la disponibilidad de una explicación mental o personal en cualquier caso en el que un agente no tenga la intención correspondiente. Tampoco importa que sólo contemos con ideas esquemáticas y anecdóticas sobre la manera de completar los detalles de toda la gama de explicaciones esbozadas, pues lo mismo se aplica a muchos de nuestros conceptos disposicionales ordinarios.

Pero, ¿por qué en todos esos casos *tiene* que haber una explicación mental o personal del fracaso del deseo más fuerte del agente para involucrar su voluntad? Quizás el «signo primitivo» del querer sea «tratar de obtener» (Anscombe 1963: sección 36). Pero un signo, por primitivo que sea, no es más que eso: no se trata de un concomitante inevitable, ni siquiera de un concomitante normal, de aquello de que es signo. Y resulta difícil creer que admitir, en contra de P1*, la posibilidad de algunos casos «inexplicables» —es decir, casos para los cuales no hay ninguna explicación de la falta de intención del agente en términos de otras características de su vida mental o de su personalidad— pueda reducir al

sinsentido las atribuciones de deseos a los agentes. Desde luego, siempre que a un agente le falte la intención apropiada puede hablarse, si se desea, de su «inercia»: pero ello no proporciona ni siquiera el esquema de una explicación del tipo que sea.

Se obtiene una defensa tal vez más convincente de un principio como P1* cuando se trata de conectar ese principio con la ortodoxia vigente sobre la teoría causal de las razones de la acción. Para nuestros propósitos actuales, esa ortodoxia, o por lo menos la parte de ella que aquí es pertinente, puede entenderse como sigue: si un agente actúa con acierto por alguna razón que tiene para actuar así, entonces (a) esa razón consiste en una combinación de deseos y creencias tal que, en virtud del contenido proposicional de esos estados mentales, su acción queda *racionalizada*, se vuelve al menos mínimamente razonable; y (b) esa razón es la *causa* de su acción. (Nos basamos tácitamente en la condición (a) al inicio de esta discusión sobre el deseo y la acción.) Ahora bien, supongamos que un filósofo queda convencido por los argumentos conocidos en favor de esta teoría causal de las razones de la acción, y supongamos también que acepta un argumento conectivo que pretende mostrar que el rechazo de todas las concepciones del deseo como poder activo exige el rechazo de la teoría causal de las razones de la acción. Así pues, obviamente este filósofo se sentirá obligado a rechazar el rechazo de todas las concepciones del deseo como poder activo.

Uno de estos argumentos conectivos podría ser el siguiente. Cualquier afirmación causal singular verdadera implica la existencia de una ley causal que la cubra. De suerte que si aceptamos, simplificando un poco, que la acción intencional de un agente es efecto de un deseo suyo, nos comprometeremos con la existencia de una ley causal correspondiente que cubra el caso. Pero el rechazo de la concepción del deseo como poder activo, la aceptación de casos «sin explicación», consiste precisamente en el rechazo de la tesis de que debe haber una ley semejante que conecte deseos de esa clase con la intención, ya no digamos con la acción. El sentido de ese rechazo, repito, no es que tengamos solamente ideas muy esquemáticas acerca de cómo articular las condiciones mentales y personales bajo las que se producirán la intención y la acción, sino que, más bien, ese rechazo niega que tenga que haber un conjunto semejante de condiciones que haya que articular.

Es innecesario detenerse en el supuesto de la naturaleza determinista de las leyes causales, por tendencioso que pueda ser, ya que Davidson nos ha armado con una replica más directa al argumento conectivo. El

enunciado correcto de la relación entre afirmaciones causales singulares y leyes causales «de cobertura» es el siguiente:

> si «*a* causó *b*» es verdadero, entonces hay descripciones de *a* y de *b* tales que el resultado de sustituirlas por «*a*» y «*b*» en «*a* causó *b*» está implicado por [una ley junto con un enunciado de que se produjo un caso particular de una clase especificada a la que *a* pertenece]; y lo inverso también se sostiene si se imponen restricciones adecuadas a las descripciones. (Davidson 1967: 159-160)

Así, que un deseo sea la causa de una acción no implica que exista una ley expresada en términos de deseos de esa clase. De modo comparable, entonces, el principio P1* no obtiene apoyo de la consideración de que cuando un agente en verdad intenta hacer lo que más quiere, su intención es efecto de su deseo. La debida apreciación del fundamento nomológico (presunto requisito) de esa relación causal puede requerir una modificación radical de nuestros términos de referencia. La necesidad de colocar al deseo que da como resultado una acción o incluso una mera intención dentro del nexo causal no exige la aceptación del deseo en general como un poder activo.

Consideremos a dos agentes con personalidades idénticas, que hasta ahora han disfrutado de vidas mentales idénticas. Imaginemos incluso que las «realizaciones físicas» de sus vidas mentales y de sus rasgos de personalidad han sido siempre y siguen siendo idénticas. Pero ahora uno de ellos llega a tener la intención de hacer lo que más quiere, y lo hace, y el otro no. Entonces, presuponiendo el determinismo causal en el nivel físico, tiene que haber existido alguna diferencia física previa entre los dos agentes. Pero, ¿por qué esa diferencia física *tiene que* ser la realización o de una diferencia inadvertida entre sus vidas mentales o personales, o de una diferencia en sus capacidades? ¿Por qué no puede tratarse de una mera diferencia física? Y el abandono del supuesto de la identidad de las realizaciones físicas de lo mental y lo personal revela todavía con mayor claridad que las maniobras defensivas de que disponen los teóricos del deseo como poder activo no son más que la expresión de un dogma.

1.4 Deseo e intención

Debe subrayarse que la defensa que acabamos de ofrecer del rechazo de las concepciones del deseo como poder activo es una defensa sumamente concesiva. Tanto el supuesto de la naturaleza determinista de las leyes causales como la necesidad de respaldar un enunciado causal singular

con una ley causal que lo cubra se han aceptado por mor del argumento. Lo mismo es cierto —y esto es quizás más importante— respecto de la ortodoxia actual de la teoría causal de las razones para la acción. Quisiera ahora enfatizar brevemente la naturaleza no conclusiva de uno de los argumentos que hoy en día nos son familiares en favor de la teoría causal, pues ese carácter no conclusivo se enlaza con una idea importante que acabamos de tocar.

Podría parecer que el argumento en cuestión equivale a una defensa directa de un principio semejante a P1*. Este argumento parte del hecho de que muchas veces las personas saben lo que van a hacer antes de hacerlo, o tienen creencias razonables acerca de ello: tienen conocimiento o creencias razonables acerca del futuro en lo que concierne a sus propias acciones. En ocasiones puede tratarse de un conocimiento francamente inductivo: sabiendo que todos los días durante los últimos años me he tomado un whisky puro a las once en punto de la noche, tengo justificación, *ceteris paribus*, para creer que procederé de la misma manera esta noche. Ese conocimiento o esa creencia razonable acerca de mis acciones futuras es, por lo menos en principio, igualmente accesible para otra persona. Pero supongamos que, molesto porque acabo de descubrir lo predecible que soy, llego a tener un fuerte deseo de cambiar esta noche el whisky por un *tequila sunrise*; entonces puedo saber, o creer razonablemente, que voy a beber un *tequila sunrise* esta noche a las once en punto. Pero, ¿*cómo* puedo saberlo, o tener una creencia razonable al respecto? Una idea (véase Pears 1964) es que, so pena de aceptar un misterioso acto de premonición, tiene que darse el caso de que: (i) sé algo acerca del presente y (ii) sé, o creo razonablemente, que el objeto de ese conocimiento presente es la clase de cosa que causa futuros hechos de beber *tequila sunrise*. Es decir, continúa la idea citada, conozco mi deseo presente y sé que ese deseo es la clase de cosa que causa futuros hechos de beber *tequila sunrise*. Está claro que en un caso dado puede resultar que me equivoque: más tarde podría cambiar de opinión, podría no haber tequila que beber, podría morir antes de la hora feliz. Aun así, la idea estriba en que, a menos que conozca mi deseo actual junto con una tosca generalización causal similar a P1* que enlace deseos de ese tipo con tipos de acción futura, mi conocimiento o creencia razonable respecto de lo que voy a hacer resultaría totalmente misterioso.

Ese argumento merece un tratamiento más paciente que el que va a recibir aquí. Su réplica obvia toma la forma de una explicación alternativa de nuestro conocimiento sobre nuestras acciones futuras (véase

O'Shaughnessy 1980). En términos muy esquemáticos, una explicación semejante tiene la siguiente estructura: primero, se introduce una noción de *conocimiento sin observación* en el contexto del conocimiento de una persona acerca de sus acciones presentes —un conocimiento práctico acerca de lo que se está haciendo intencionalmente en este momento (véase Anscombe 1962: sección 8)—. Está claro que uno puede equivocarse acerca de lo que hace intencionalmente en este momento: pero, en tal caso, la falta se encuentra, por así decirlo, en la acción y no en la creencia. Y segundo, la misma noción de conocimiento práctico se amplía para poder aplicarla a acciones futuras: así como ahora sé lo que estoy haciendo intencionalmente, también sé ahora lo que haré intencionalmente en el futuro. Desde luego, cualquier pretensión semejante de conocimiento sobre las acciones futuras exige una cláusula *ceteris paribus*; y, por supuesto, aun cuando las cosas no varíen, podría haber algún error práctico en mi futura acción conforme a la intención. Pero, en el último caso, la falta vuelve a encontrarse en la acción, no en la creencia.

Ciertamente, esa explicación alternativa se ha expuesto en términos sumamente esquemáticos; pero la forma de exponerla sirve para llamar la atención sobre otro defecto del argumento ofrecido en favor de la teoría causal —un defecto que vuelve innecesario, para nuestros propósitos actuales, abandonar el nivel esquemático a fin de tomar, con minucioso detalle, una resolución entre ambas explicaciones acerca del conocimiento que tiene una persona sobre sus acciones futuras—. La explicación alternativa que se ha expuesto partía de nuestro conocimiento práctico sobre lo que estamos haciendo intencionalmente en este momento y trataba de ampliar esa noción a nuestro conocimiento presente acerca de lo que en el futuro haremos intencionalmente. La idea aquí es obvia: en este argumento en favor de la teoría causal, se pasó de la estructura general de la explicación del conocimiento futuro de la acción —los elementos (i) y (ii)— a la afirmación de que el deseo constituye el factor apropiado en la realización más concreta de esa estructura. Pero está muy poco claro que el deseo sea el factor apropiado; dentro de los términos de esa explicación, la decisión o la intención serían candidatos mucho más convincentes. Una vez más, parecería que un teórico borra la distinción crucial entre el deseo, por un lado, y la intención, digamos, por el otro. Es posible que una resolución detallada en favor de una u otra explicación del conocimiento de la acción futura pusiera de manifiesto la necesidad de ciertas generalizaciones causales toscas que vinculen las intenciones o decisiones del presente con las acciones futuras; pero eso dejaría sin tratar la cuestión del deseo.

1.5 Explicación de la acción

En la búsqueda de argumentos en favor de un principio como P1*, dejemos la causación por el tema distinto, aunque relacionado, de la explicación. El rechazo del principio P1* equivale a aceptar la posibilidad de que la totalidad de las otras verdades mentales y personales acerca de un agente sea tan compatible con que no tenga la intención de hacer lo que más quiere hacer como lo es con que tenga esa intención y actúe en consecuencia. Pero si ello fuera posible, ¿cómo podría ser que la referencia a esa totalidad de otras verdades mentales y personales, ya no digamos a una porción favorecida de la misma, pudiera servir para *explicar* la intención y la acción del agente en los casos en que se producen?

La suposición de que el valor de las explicaciones que hacen referencia a la vida mental y personal de los agentes depende de la verdad de principios generales de la forma de P1*, se nutre de muchas fuentes. Todo lo que quiero hacer aquí es una breve observación diagnóstica destinada a exponer una de esas fuentes, un prejuicio pertinente que es natural para el hombre. Si se comienza pensando que la vida mental y personal de otro se halla necesariamente fuera de la propia mirada perceptual —es decir, si se supone que cualquier atribución defendible de mentalidad o personalidad al otro tiene que ser resultado de una inferencia—, se terminará naturalmente por pensar en atribuciones defendibles de estados mentales y rasgos de personalidad a los otros conforme al modelo hipotético-deductivo de inferencia explicativa de lo que no se percibe (por imperceptible). Entonces, una vez convencidos de la rectitud de ese modelo, parecerá simplemente obvia la necesidad de principios generales de la forma de P1 y P1* para reforzar nuestras especulaciones sobre los mecanismos hipotéticos.

Quizás el punto de arranque de esas reflexiones nos resulte completamente natural. Pero lo natural no es siempre inevitable ni siempre correcto. Ese punto de arranque, con su concepción cartesiana de la mentalidad y la personalidad como lo interno fugitivo, puede rechazarse sin que por ello se necesite abrazar por toda cosa una concepción conductista de lo externo. Este doble rechazo aparece abierto ante nosotros al tomar literalmente la idea de que, por ejemplo, podemos ver el deseo del seductor en su rostro. Su deseo no queda reducido a una nada al rechazar el pensamiento de que se trata de un mero algo cuya presencia y naturaleza sólo es accesible a los demás por inferencia; puede ser *esa cosa* que los otros *ven* (véase Wittgenstein 1953: sección 304). No por estar equipados con el lenguaje de la mente estamos en aptitud de sostener una nueva

gama de hipótesis explicativas sobre lo perceptualmente inaccesible, y tampoco nos iniciamos en una teoría nueva que podamos aducir luego en la interpretación inferencial de lo perceptualmente accesible; más bien, con ello hemos puesto dentro de nuestro alcance perceptual inmediato, no inferencial, un nuevo dominio de hechos. Y esta perspectiva sobre el conocimiento de la vida mental y personal de los otros no se ve más socavada por los hechos del error ocasional y la inferencia ocasional que la perspectiva comparable sobre nuestro conocimiento del mundo exterior.

1.6 Caracterizaciones de deseabilidad

Expresando la «opinión irreflexiva corriente», Russell escribió:

> Pensamos que el contenido del deseo es exactamente igual que el contenido de una creencia, mientras que la actitud que se asume hacia el contenido es diferente. De acuerdo con esta teoría, cuando decimos: «quiero que llueva» o «espero que llueva», expresamos, en el primer caso, un deseo, y en el segundo, una creencia, con un contenido idéntico, a saber, la imagen de la lluvia. Sería fácil decir que tal como la creencia es una clase de sentimiento en relación con este contenido, así el deseo es otra clase. (Russell 1921: 58)

Ya expresamos nuestras dudas acerca de la atribución que hace Russell a la concepción del deseo que tiene el sentido común de algún tipo de papel universal para las imágenes y los sentimientos; en vez de ello —se dijo—, tal concepción consiste en que el deseo es una actitud proposicional que se dirige a un contenido proposicional específico. Aun así, parece claro de todas maneras que Russell habría rechazado esta última concepción.

Tal rechazo colocaría a Russell en franca oposición no sólo a la concepción de que el deseo es una actitud proposicional, sino también a otra concepción del deseo más evasiva que en ocasiones ha acompañado a la primera. Ya nos hemos topado (p. 33) con la doctrina tomista de que, en la medida en que una persona actúa intencionalmente, actúa a la luz de un bien imaginado. Esa doctrina sobre la naturaleza de la acción intencional invita a otra sobre la forma de entender las acciones intencionales como acciones intencionales, a saber: que la racionalización de una acción intencional dada debería revelar el bien imaginado a cuya luz el agente en cuestión llevó a cabo esa acción. Simplificando para alejarnos de algunas complejidades muy importantes pero que no son pertinentes aquí, y apoyándonos en la ortodoxia contemporánea sobre la estructura general de tales racionalizaciones, podríamos suponer que ese requisito se satisface con una especificación adecuada del contenido de cualquier deseo que se

aduzca dentro de una racionalización plenamente articulada de la acción intencional en cuestión.

No todos los deseos provocan una acción, ni todas las atribuciones de deseos se hacen como parte de la elaboración de racionalizaciones de acciones intencionales. Sin embargo, resulta difícil negar que la aptitud de las atribuciones de deseos para desempeñar en las racionalizaciones de las acciones intencionales el papel esbozado, es esencial para la cuestión del concepto de deseo. Siendo esto así, lo anterior sugiere que se adopte la siguiente condición general para las atribuciones de deseos: una atribución explícitamente aceptable de un deseo específico a un agente es aquella que, en virtud de la especificación del contenido del deseo, sirve para revelar el bien imaginado a cuya luz el agente actuaría si realmente actuara conforme a ese deseo.

¿En qué consistiría «revelar el bien imaginado» mediante la especificación del contenido de un deseo dado? G.E.M. Anscombe afirmó en cierta oportunidad que sería un «franco disparate» decir: «Los filósofos han enseñado que cualquier cosa puede ser un objeto de deseo, así que puede no ser necesario que caracterice estos objetos como deseables en algún sentido; simplemente ocurre que los quiero» (Anscombe 1963: sección 37). Esto podría sugerir la siguiente reformulación de la condición para las atribuciones de deseo explícitamente aceptables que se extrajo de la doctrina de Santo Tomás: una condición necesaria para la aceptabilidad de una atribución de un deseo específico a un agente consiste en que quien hace la atribución haya «alcanzado y vuelto inteligible» una «caracterización de deseabilidad» del estado de cosas (potencial) ahí especificado. Esa caracterización de deseabilidad revelará directamente cuál es el bien imaginado a cuya luz el agente actúa si en efecto actúa conforme a ese deseo. Por lo tanto, o bien el contenido proposicional especificado en la atribución de deseo debe ser una caracterización de deseabilidad, o bien, por así decirlo, debe estar respaldado por una caracterización semejante. Está claro que quien, como Russell, niega la categoría del deseo como actitud proposicional, no podrá encontrar ningún sentido en esta otro idea que pretende enlazar la noción de deseo con la de una caracterización de deseabilidad.

Más adelante habrá que elucidar y matizar considerablemente esta otra idea. Por el momento puede bastar con que se agregue lo siguiente: se ha alcanzado una caracterización de deseabilidad del objeto de deseo cuando el cuestionamiento en torno a por qué el agente quiere ese objeto se lleva a buen término. La caracterización de deseabilidad es la que resulta

adecuada para comunicar la deseabilidad del objeto: una vez que se ha dado, o alcanzado, esa caracterización, el bien imaginado se ha manifestado ante nuestros ojos.

1.7 «Deseos animales»

Podemos apreciar ahora, según creo, algo de la complejidad de la afirmación 4 de Russell, de que el estudio de los animales no humanos es en muchos sentidos la mejor preparación para el análisis del deseo. No obstante, comencemos por un caso que en apariencia es casi humano.

> Sísifo, como se recordará, reveló secretos divinos a los mortales, y por esto los dioses lo condenaron a empujar una roca hasta la cima de un monte, de donde la roca caía rodando de inmediato para que Sísifo volviera a empujarla hasta la cima, de donde volvía a caer, y así sucesivamente una y otra vez, *por siempre*. (Taylor 1984: 256–257)

Según Richard Taylor, esto es «una clara imagen de la existencia sin sentido». Luego, usando el siguiente ejemplo, Taylor introduce su explicación respecto de cómo puede eliminarse ese sinsentido:

> [Supongamos que a los dioses] luego se les ocurrió ponerse perversamente clementes imponiéndole [a Sísifo] el extraño e irracional impulso [...] de empujar rocas [...] [E]n términos más gráficos, supongamos que consiguieron esto implantándole una sustancia que afectaba de ese modo su carácter y sus instintos [...] Usando este mecanismo, lograron darle a Sísifo precisamente lo que éste quería: lo hicieron desear justamente lo que le habían infligido. (*Ibid.*: 259)

En líneas más generales, Taylor concluye que podemos

> reintroducir lo que se ha [...] dejado firmemente de lado en un esfuerzo por ver con objetividad nuestras vidas y la existencia humana; a saber, nuestra propia voluntad, nuestro profundo interés en lo que estamos haciendo. (*Ibid.*: 266)
>
> El sentido de la vida procede de nuestro interior, no se nos dispensa desde afuera, y tanto por su belleza como por su permanencia excede con mucho cualquier cielo que los hombres hayan soñado o anhelado jamás. (*Ibid.*: 268)

Está claro que Taylor supone que ha descrito un caso en el que los dioses inducen realmente «dentro de» Sísifo el deseo de empujar rocas. Y sus

expresiones concomitantes acerca de un «extraño e irracional impulso» y de «instintos» sugieren por qué supone que ha hecho justamente eso. Pues tales expresiones, al igual que sus conclusiones generales, indican que Taylor suscribe tácitamente una concepción del deseo como poder activo, como mera disposición a actuar, desculdando por completo las consideraciones que se siguen de la aceptación de la doctrina de Santo Tomás. Si se ha inducido en Sísifo una disposición a empujar rocas, éste posee entonces, sin mayores problemas, el *deseo* de empujar rocas.

¿Para qué diablos, o demonios, querría el Sísifo de Taylor empujar rocas? ¿Qué ve en ello? El caso sería diferente si se preocupara por aplacar la furia de los dioses con la esperanza de evitar mayores castigos o de acabar con el actual. Sería también diferente e incluso más interesante si le preocupara la justicia de su castigo y estuviera convencido de ella. (Tal vez perdone la forma de la pena en vista de los limitados recursos financieros de los dioses.) En ambos casos podríamos comprender lo que se encuentra «dentro» de Sísifo en función de su percepción de lo que se halla «fuera»; pero, en el ejemplo descrito por Taylor, la única manera en que podemos entender el supuesto deseo alojado «dentro» de Sísifo es en términos de algo que está literalmente dentro de él: la sustancia implantada por los dioses; y esto no nos da ninguna idea en absoluto sobre ningún bien imaginado a cuya luz se supone que Sísifo actúa.

Es posible, sin duda, inducir impulsos e instintos de los tipos más insólitos en los seres humanos —y en los loros, las ratas y los moluscos marinos—. Pero no deberíamos apresurarnos tanto a servirnos de las nociones de deseo y acción intencional al describir los resultados de tales maquinaciones. Comparemos el Sísifo de Taylor con el caso imaginado en que a Sísifo le preocupa aplacar la furia de los dioses con la esperanza de evitar mayores castigos o de acabar con el actual. La analogía que invita a ampliar al Sísifo de Taylor el discurso sobre el deseo es clara: en ambos casos, el «comportamiento» de Sísifo es el mismo. Pero también hay diferencias cruciales. El Sísifo de mi ejemplo bien puede obtener su deseo de empujar rocas, o puede verse llevado a modificar la fuerza de ese deseo, por lo que anteriormente (p. 32) denominé «deliberación protopráctica» encaminada a responder la pregunta: «¿Qué es lo que más quiero hacer?». Su deseo y la fuerza del mismo pueden ser consecuentes con otras creencias, deseos y concepciones del mundo que él sostiene. Aun más, en mi ejemplo, el deseo de Sísifo no es una mera disposición a «comportarse» de un modo rutinario; puede dar lugar a patrones de acción ilimitadamente variados, dependiendo de los otros deseos y creencias que tenga. Si el Sísifo de mi ejemplo pensara que también otras acciones

lograrían aplacar la furia de los dioses, bien podríamos encontrarlo intentando alguna de ellas; si el Sísifo de mi ejemplo llegara a creer que los dioses están muertos, podemos suponer que suspendería su actividad de empujar rocas; si el Sísifo de mi ejemplo empezara a dudar de su creencia de que su actividad está aplacando la furia de los dioses, y quisiera por tanto ponerla a prueba, podría hacer una pausa momentánea en su trabajo; si el Sísifo de mi ejemplo se sintiera completamente exhausto y quisiera descansar un momento, podría arreglárselas para romperse «accidentalmente» una muñeca. Para comprender cualquiera de los casos de acciones alternativas, será necesario hacer referencia a la caracterización general de deseabilidad cuya aceptación condujo a mi Sísifo, en un principio, a acciones particulares de empujar rocas: la evitación de un castigo ulterior o la terminación del actual. Y esto requerirá atribuirle a Sísifo la posesión de los *conceptos* pertinentes, como lo exigiría, desde luego, la explicación de cualquier deliberación protopráctica que hubiera emprendido mi Sísifo al obtener su deseo. Por su parte, para el Sísifo de Taylor, o para un simple molusco marino invadido por un instinto, no se requiere esa atribución conceptual. El Sísifo de Taylor no necesita tener en absoluto ninguna concepción de su «actividad»; en consecuencia, no necesita tener de ella ninguna concepción de deseabilidad. Hay aquí un eco contemporáneo de la coincidencia histórica del apego a las concepciones del deseo como poder activo y la negación tácita de la categoría de los deseos como actitudes proposicionales: una coincidencia que, coherentemente, requiere que se abandone el empleo del concepto de acción intencional.

En *Ulises y las sirenas* (1984 [1989]*), Jon Elster expresa algunas ideas interesantes acerca de «el rasgo característico del hombre» y el momento en que «la *mente* ingresa en la arena de la evolución». El hombre es capaz tanto de *esperar* como de utilizar *estrategias indirectas*; es decir, el hombre es capaz de renunciar ahora a una posibilidad favorable a fin de conseguir otra aún más favorable en el futuro, y también de aceptar ahora una posibilidad desfavorable a fin de obtener otra muy favorable posteriormente. Estas dos aptitudes son constitutivas de la capacidad para lo que Elster denomina *optimización global* (*ibid.*: 15–17). Esta capacidad también se encuentra en algunos animales no humanos, pero en ellos «el comportamiento de optimización global [...] se encuentra en situaciones sumamente específicas y estereotipadas». Mientras que, afirma Elster, «el

* Las referencias entre corchetes remiten a las versiones en castellano de las obras citadas por el autor. [N. de los tt.]

rasgo característico del hombre consiste en [...] una *capacidad generalizada de optimización global* que se aplica aun en situaciones cualitativamente nuevas». Incluso, en casos de optimización global específica para una situación, «no es preciso recurrir a estructuras intencionales o mentales»; pero «el uso de estrategias de optimización global en situaciones nuevas tiene que implicar un análisis del contexto, un sondeo de los varios movimientos posibles y, por último, la elección deliberada entre ellos». Así pues, la capacidad generalizada para la optimización global requiere la habilidad de referirse al futuro y a lo meramente posible, y Elster sugiere que «con esta capacidad generalizada, la *mente* ingresa en la lid evolutiva».

El Sísifo de mi ejemplo, una vez que su deseo de empujar rocas se ha hecho comprensible en términos de su reconocimiento de la deseabilidad de evitar mayores castigos o de terminar el presente, bien podría servir para ejemplificar, *en esta situación particular pero novedosa*, su capacidad generalizada para la optimización global. Ante la novedad, él podría declinar, por ejemplo, un ofrecimiento de ayuda en su trabajo ya que puede considerar que tal asistencia, al enardecer la furia de los dioses, podría reducir sus posibilidades futuras de libertad (espera). O podría decidirse ahora por la desagradable opción de empujar una roca más pesada, ya que puede considerar la posibilidad de que, al aplacar la furia de los dioses, incrementaría por consiguiente sus posibilidades futuras de libertad (estrategias indirectas). Por su parte, para el Sísifo de Taylor, *según se lo describe*, no parece haber ningún medio inteligible por el que se pudiera generar esa posible diversidad.

Quizás lo más extraño acerca de la concepción del deseo de Richard Taylor, como cualquier concepción que descuide las intuiciones que se siguen de la doctrina de Santo Tomás, es que sus partidarios simplemente tienen que haber pasado por alto los diferentes grados y formas en que podemos entender los objetos de muchos deseos. En una discusión sobre las opiniones de Taylor, David Wiggins proporcionó algunos buenos ejemplos:

> existe [...] una diferencia —en la que insistiremos como participantes— entre la vida de un hombre que efectúa cierta contribución a una sociedad que tiene una historia continua y una vida llevada en el nivel de un criador de cerdos sureño que [...] compra más tierra para cultivar más maíz para engordar más puercos para comprar más tierra [...] (Wiggins 1976: 100-101)

Y también:

> Al participante puede parecerle mucho más difícil explicar qué tiene de bueno comprar más tierra para cultivar más maíz para engordar más puercos

> para comprar más tierra [...] que explicar lo bueno de cavar una zanja con un hombre que a uno le agrada, o de compartir con el mismo hombre una charla o un trago mientras el sol agoniza. (*Ibid.*)

Si los deseos humanos únicamente equivalieran a impulsos extraños e irracionales, a meros instintos, resultaría completamente ininteligible que existieran esas diferencias claras, aceptadas e importantes entre las actividades de estos dos hombres. La única manera en que un deseo podría resultarnos enigmático sería como algo estadísticamente excepcional; de ese modo, me temo, no es el criador de cerdos sureño quien en la actualidad representaría un enigma.

Por último, cabría señalar que los deseos animales genuinos —no los «instintos animales» del Sísifo de Taylor y de los moluscos marinos— no son meros instintos ininteligibles: el Sísifo de Taylor es un modelo general tan malo para comprender los deseos animales como para comprender nuestros deseos. Parte del problema con la afirmación 4 de Russell, de que el estudio de los animales no humanos es en muchos sentidos la mejor preparación para el análisis del deseo, consiste en que, al combinarla con las otras afirmaciones de Russell sobre el deseo, invita a explicar los deseos animales precisamente según las directrices de Taylor.

La mayor parte de los elementos de la concepción russelliana del deseo que aquí hemos criticado se encuentran, creo, en la mayoría de las concepciones tácitas del deseo sostenidas por otros filósofos. Quizás la negación de la categoría del deseo como actitud proposicional resulta ahora comparativamente rara, y quizás la exageración del papel del sentimiento también es ahora menos frecuente; pero todavía se halla muy difundida una concepción disposicional falsa, vaya acompañada o no por el apego a una concepción causal de esa supuesta disposición. Y más de un filósofo contemporáneo persiste en la idea de que los animales no humanos constituyen un punto de partida clave para reflexionar sobre la naturaleza del deseo. Los errores de Russell no eran excentricidades individuales.

Es probable que una reacción casi universal ante el tratamiento dado aquí a la concepción de Russell sea la siguiente: si el deseo no es eso, entonces ¿qué es? Queda por ver si es posible lograr que algunas de las ideas aducidas en esta crítica de la concepción de Russell se presten a una concepción más verosímil de lo que es el deseo o, por lo menos, a un punto de vista más adecuado del lugar que ocupa el deseo en la vida mental y activa de los seres humanos.

CAPÍTULO 2

LAS DISTINCIONES DEL DESEO

> Los hombres siguen casi siempre las sendas trazadas por otros y proceden en sus asuntos por imitación, aun cuando no puedan mantenerse por entero dentro de esas sendas ni emular la proeza de sus modelos. Así, un hombre prudente debe seguir siempre las huellas de los grandes hombres e imitar a quienes se han destacado. Si sus propias proezas no llegan a compararse con las de ellos, al menos estarán envueltas en cierto aire de grandeza.
>
> NICOLÁS MAQUIAVELO

> No hay en el mundo criatura más desdichada que el fetichista que suspira por el zapato de una mujer y tiene que contentarse con la mujer entera.
>
> KARL KRAUS

2.1 NECESIDADES, CARENCIAS, DESEOS

De acuerdo con la falsa concepción clásica, un deseo es «un algo introspectivo» (un sentimiento) que constituye una disposición o tendencia a hacer algo (una fuerza que nos mueve), y que no contiene ninguna representación de ningún estado de cosas, ni real ni meramente imaginario. Pero, por natural que pueda ser esa concepción del deseo, tenemos motivos para pensar que es del todo errónea. Sin embargo, evitando las falsificaciones de los «sentimientos» y la vaciedad de las «disposiciones», ¿puede encontrarse una concepción más verídica de lo que el deseo es?

Es posible arribar a una respuesta más instructiva (véanse las pp. 54–62), aunque en última instancia equivocada, mediante una consideración de las distintas nociones de *necesidades* y *carencias*. La noción más general de la necesidad de algo se aplica cuando ese algo es necesario, se necesita, para la realización de cierta situación (véase Wiggins 1985). El término *«needcessity»*,[1] que se usó en el sur de Estados Unidos hasta, por lo menos, fines del siglo pasado, representa una resurrección tan placentera

1. Combinación de las palabras *need* y *necessity*, que significan, ambas, «necesidad». [N. de los tt.]

de este pensamiento aristotélico, como disonante es la expresión que de él hace. Un modismo menos chocante al oído, que capta la misma idea de una necesidad surgida de los hechos del caso, es *«it is needful that»*.[2] Hay alguna razón para pensar que esa noción más general de necesidad se deriva de otra noción un poco más estrecha, a saber, la de una necesidad básica que se aplica cuando el algo necesitado es necesario para el florecimiento y el bienestar del sujeto al que se atribuye la necesidad. Pues, sin duda, tendríamos que sentirnos impresionados por el hecho de que las atribuciones de necesidades que conllevan esa relativización más específica, aunque tal vez tácita, son en general inteligibles de manera independiente, mientras que las atribuciones basadas en la noción enteramente general de una necesidad, requieren, para resultar inteligibles, una guía contextual específica respecto de cuál es la relativización apropiada.

La noción distintiva de carencia es captada por quienes aprecian debidamente la primera línea del Salmo 23: *«The Lord's my shepherd, I'll not want»*.[3] Esa versión se encuentra, entre otros lugares, en los *Scottish Metrical Psalms* de 1650;[4] el *Prayer Book*[5] de 1662 explica las cosas, aunque a un precio poético elevado: *«The Lord is my shepherd: therefore I can lack nothing»*.[6] Una carencia es una falta: mi carencia ahora es la inteligencia, la de los lectores es probablemente la paciencia. (El *Dictionary of the English Language* de Samuel Johnson define *wantwit*[7] como tonto o idiota.) No es de sorprender que exista la correspondiente forma verbal distintiva, aunque el común descuido en el uso de las preposiciones como partes de los verbos puede conducir a una confusión entre esa forma y el verbo que se emplea más naturalmente al hablar de los deseos.[8] Así pues, yo carezco de [*want for*] inteligencia; el lector de paciencia. Alexander Pope estaba jugando con esta confusión potencial cuando escribió:

> «Con tanta gracia, con tanta sabiduría,
> Decidme, ¿qué puede querer Cloe?» — Quiere un corazón.[9]

2. Correspondería a algo como «es necesitado que». [N. de los tt.]
3. «El Señor es mi pastor, nada me faltará.»
4. Una versión métrica de los salmos publicada en Escocia en 1650. [N. de los tt.]
5. *Libro de Oraciones*. [N. de los tt.]
6. «El Señor es mi pastor: por lo tanto, nada puede hacerme falta.»
7. Literalmente, sería algo como «ingeniocarente». [N. de los tt.]
8. En inglés, las preposiciones modifican el significado de los verbos: *to want* es «querer», mientras que «to want for» significa «carecer de». [N. de los tt.]
9. *'With ev'ry pleasing, ev'ry prudent part, / Say, what can Cloe want?' – She wants a heart.* Donde «querer» [*want*] puede intercambiarse por «desear» o «carecer de». [N. de los tt.]

Ni cualquier cosa ausente ni cualquier laguna constituye una carencia o una falta. La desnutrición, por ejemplo, es algo que me falta pero que no necesito. En realidad, esto sugiere una fórmula limpia sobre cuándo una ausencia equivale a una carencia: cuando y sólo cuando lo ausente es necesario. (De ahí la agudeza de la reformulación del Salmo 23 en *Private Eye* (11 de noviembre de 1988): «La Señora es mi pastora, no me veré privado de ningún bien de subsistencia necesario».) Al considerar esa propuesta, nos asombrará el paralelismo que se da entre las atribuciones de carencias o faltas y las atribuciones de necesidades en cuanto a los variados papeles que cumplen las consideraciones contextuales específicas, incluidas las relativizaciones particulares, para volver inteligibles esas atribuciones. Así, por ejemplo, las consideraciones cuya apreciación sería necesaria en un caso particular a fin de que la idea de que necesito acatarrarme tuviera sentido, también tendrían que aducirse a fin de conferirle un grado similar de inteligibilidad, en esa misma situación, a la idea de que carezco de catarro. La diferencia entre el concepto de una necesidad y el de una carencia podría parecer que se convierte entonces en la idea anodina de que, aunque puedo necesitar lo que de hecho tengo, no puedo, sin embargo, carecer de ello.

Pueden hacerse atribuciones inteligibles de necesidades y carencias a cosas que no son capaces de actuar intencionalmente, e incluso a cosas que no tienen vida mental alguna. Aun cuando el sujeto de tales atribuciones sea un ser humano normal, no es preciso que la atribución implique nada de la vida mental del sujeto. La carencia de corazón de Cloe dice mucho sobre su vida mental, reconozca ella o no la verdad de lo que se expresa. Pero esto es así únicamente porque se trata de un caso en que la falta en cuestión es de índole espiritual. Con que se elimine el aspecto metafórico y se disponga de algún otro medio distinto y eficiente para bombear la sangre a través de su cuerpo, nada se seguirá acerca del estado mental de Cloe del hecho de que le falte el corazón.

Lo anterior sirve para poner en duda cualquier idea de que las autoatribuciones de carencias y necesidades se fundan en general en una «relación epistemológica» distintiva, en la que se considera que cada sujeto se encuentra con su propio estado mental. Por el contrario, se da lugar al pensamiento de que la capacidad de autoatribuirse carencias y necesidades puede reflejar meramente la sensibilidad del sujeto a las relaciones que se dan entre sus circunstancias actuales y su bienestar (potencial o real). Los fundamentos de esas autoatribuciones de carencias y necesidades pueden no diferenciarse de los de atribuciones comparables referidas

a otros. Así pues, queda abierta la posibilidad de que los otros estén en una mejor posición para juzgar sobre las necesidades y carencias del sujeto que el sujeto mismo.

Consideremos las autoatribuciones de carencias o faltas específicamente mentales. En tales casos está implicado, en parte, el presunto reconocimiento de una ausencia mental, de una laguna en la vida mental. Sería evidentemente un camino erróneo sostener que tales ausencias deben anunciarse. No es necesario recurrir a la parafernalia del autoengaño, el subconsciente y demás para entender la posibilidad de que Cloe no tome conciencia de una verdad central sobre su vida mental: su falta de corazón. Además, incluso si Cloe, después de muchas búsquedas introspectivas, se atribuye una falta de corazón, no hay ninguna razón general para pensar que su afirmación es incuestionable. La indagación del alma no necesita producir un inventario exhaustivo del contenido del alma.

El reconocimiento, verdadero o falso, de un agente de alguna necesidad o carencia propia puede dar origen a un deseo correspondiente. Y la perplejidad que sintamos acerca de algún deseo particular del agente puede eliminarse contemplando cómo surge ese deseo del reconocimiento verdadero o falso de una necesidad o carencia correspondiente.

Deben presentarse inmediatamente dos reparos. No estoy sosteniendo que las necesidades o carencias sólo provoquen deseos a través del reconocimiento del agente de esas necesidades y carencias. No parece imposible, por ejemplo, que un agente se vea conducido a apreciar cierta necesidad o carencia suya a través de una reflexión sobre un deseo que tiene y que ha surgido de esa necesidad o carencia. Tampoco estoy sosteniendo que todos los deseos puedan entenderse como deseos que surgen de las necesidades o carencias del agente. Tal punto de vista sólo puede parecer defendible cuando se desdeña la variedad de los deseos humanos, o cuando se postulan voluntariosamente *ad hoc* las necesidades y carencias de los hombres.

Muchas necesidades y carencias surgen solamente como consecuencia de deseos particulares de los agentes, deseos que pueden tener que aducirse al elucidar la afirmación de que los agentes poseen esas necesidades y carencias. Pero las necesidades y carencias básicas, las requeridas para el florecimiento y el bienestar del sujeto, no presentan esa relación de consecuencia con los deseos del sujeto. Por ello, tales necesidades y carencias pueden atribuirse inteligiblemente a cosas que carecen por completo de vida mental.

Aunque la pregunta sobre cuáles son exactamente las necesidades y carencias básicas de un individuo dado puede ser tendenciosa en muchas áreas, sigue siendo cierto, no obstante, que las afirmaciones que se hacen para responderla pretenden tener un carácter objetivo. Esto no se debe a que tales preguntas sean el dominio de alguna investigación científica supuestamente libre de valorizaciones. Sin duda, la consideración del *tipo* de cosa que está ejemplificado por un individuo dado –planta, ser humano, tigre de Bengala– puede revelar algunas de las necesidades y carencias básicas de ese individuo por medio de las investigaciones científicas apropiadas. Pero pueden suscitarse discusiones formidables acerca de lo que constituye exactamente, digamos, el florecimiento y el bienestar de un ser humano individual; y es propio del peor tipo de cientificismo miope insistir en que tales disputas sólo se resuelven de manera adecuada a través de nuevas investigaciones científicas libres de valorizaciones. Cualquier explicación total semejante del asunto carecerá en efecto de valor –aunque no en el sentido preferido por los proponentes de esas explicaciones–. Más bien, una vez que dejemos atrás la consideración de las condiciones para la mera supervivencia, nos veremos de inmediato inmersos en cuestiones de valor, cuestiones que sólo pueden resolverse mediante el empleo de recursos distintivos como la imaginación y la empatía. Al considerar cuestiones como la de qué sería llevar una vida como *ésa*, podemos aprender mucho más, por ejemplo, de las creaciones de un novelista que de la verborrea de un científico natural. Sin embargo, pese a todas las dificultades y complejidades que se introducen con ello, la objetividad no queda en modo alguno inmediatamente proscrita, pues ninguna de esas dificultades y complejidades debe cegarnos ante el amplio acuerdo existente en estos temas, aun tratándose de seres tan complicados como nosotros, ni ante los muchos recursos y procedimientos aceptados de que disponen quienes intervienen en disputas sobre estos asuntos.

En realidad, si la introducción de cuestiones de valor presagiara la proscripción de la objetividad, ni siquiera las necesidades y las carencias básicas para la supervivencia, descubiertas por las investigaciones científicas, podrían eludir el estigma de la subjetividad. Pues ese descubrimiento se basa, de modo crucial, en ideas de poder explicativo, calidad de la concordancia con los datos, coherencia, amplitud, simplicidad funcional, grado de comprobabilidad, concordancia con otras teorías aceptadas y eficacia instrumental, todas las cuales son en sí mismas nociones de valor (véanse Nozick 1981: 483; Putnam 1981: 127–137). Éstas guían la práctica de los científicos, figuran en las explicaciones que los mismos científicos ofrecen de sus razones para actuar como lo hacen –y, lo que es más, son materia de

disputas tan tendenciosas como las que se suscitan una vez que dejamos atrás la consideración de las necesidades y carencias básicas para la mera supervivencia—. No es solamente que una explicación científica y sin valorizaciones de las necesidades y las carencias básicas no tendría la amplitud debida; es más bien que no disponemos de una explicación semejante.

2.2 Caracterizaciones de deseabilidad y percepciones de deseabilidad

En el primer capítulo (pp. 40–42) introduje la noción de caracterización de deseabilidad en los términos siguientes: una caracterización de deseabilidad del objeto de un deseo es la que es adecuada para la comunicación de la deseabilidad de ese objeto. Ahora bien, un paradigma de una caracterización de deseabilidad es una especificación del objeto del deseo en términos que revelan que ese objeto es apropiado para satisfacer cierta necesidad o carencia que el propio agente reconoce que tiene. Tal especificación pone fin a la interrogante de por qué el agente posee ese deseo: en tales casos, estamos de acuerdo en que la cuestión ha sido respondida. Este tipo de caracterización del objeto del deseo es, en efecto, adecuada para la comunicación de la deseabilidad de ese objeto; una vez que se da esa caracterización, el bien imaginado se ha revelado ante nuestros ojos.

Esos rasgos distintivos de una caracterización de deseabilidad formulada en términos del reconocimiento, por parte del agente, de alguna necesidad o carencia propia son, creo, rasgos completamente generales de las caracterizaciones de deseabilidad. Pero permítaseme agregar que tal caracterización de deseabilidad quizá sólo pueda ser alcanzada y vuelta inteligible mediante un esfuerzo considerable. No figura dentro de mi concepción que haya que tratar a otro como un «simplón torpe y balbuciente» sólo porque no entendemos de inmediato sus deseos conforme a nuestras propias caracterizaciones de deseabilidad previamente aceptadas. No cabe duda de que el esfuerzo genuino por llegar a entender a otros podría ampliar la gama de caracterizaciones de deseabilidad que nosotros mismos aceptamos, tanto como podría profundizar nuestra comprensión de las caracterizaciones de deseabilidad que de todas maneras ya aceptábamos. Podríamos ser llevados, por ejemplo, a reconocer alguna necesidad o carencia humana que antes habíamos pasado por alto, o podríamos ser llevados a una apreciación más profunda de una necesidad o carencia que ya reconocíamos. En efecto, no hay nada que pueda descartar *a priori* la posibilidad de que el esfuerzo por comprender los deseos de los demás y sus esfuerzos recíprocos por comprender los nuestros

nos conduzca a reconocer que por lo menos algunas de nuestras caracterizaciones de deseabilidad ya aceptadas servían para distinguir bienes *merramente* imaginados.

En el primer capítulo (p. 41) le adjudiqué a Anscombe la siguiente condición para la atribución de deseos: una condición necesaria para que la atribución de un deseo específico a un agente sea aceptable consiste en que quien lo atribuye haya «alcanzado y vuelto inteligible» cierta «caracterización de deseabilidad» del estado de cosas (potencial) especificado por el contenido del deseo así atribuido. La elucidación de esa condición ha sido formulada hasta ahora en términos de atribuciones de deseos a otros; ahora tenemos que considerar las autoatribuciones de deseos.

Consideremos el caso de un extranjero entregado al proyecto de tratar de alcanzar cierta comprensión de los deseos de las personas de una cultura extraña. Tras introducirse en sus prácticas, podría llegar al punto de sentirse movido, independientemente del proyecto que adoptó, a participar en alguna actividad ritual —roer los huesos de los muertos, pulir el automóvil el domingo por la mañana—, sin poder obtener, no obstante, ninguna caracterización de esa actividad que le parezca adecuada para revelar, para captar, su deseabilidad. Él considera que la actividad ritual es «de algún modo deseable»; se ha infectado, por así decirlo, con los deseos de los miembros de la comunidad que ahora ya no le es tan extraña. Pero esas percepciones de deseabilidad, esos deseos, preceden a la consecución de formas de expresión adecuadas para comunicar dichas percepciones como percepciones de deseabilidad específicas de los objetos de esos deseos (hablar de que un hombre se ha «infectado con» un deseo puede suscitar la pregunta de qué diferencias hay entre estos casos y el Sísifo de Taylor (capítulo 1, pp. 42–46); se responde a esto sucintamente más adelante (pp. 58–59)).

Ciertamente, tal postura representará un punto de transición en la búsqueda de comprensión de ese investigador. Pero no me parece un «franco disparate» pensar que se ha alcanzado cierto nivel de comprensión. Nuestro investigador estará ahora tan perplejo ante sí mismo como lo está ante quienes lo rodean; pero esa perplejidad tendrá un carácter muy distinto de la que experimentó en su primer encuentro con las actividades de la comunidad extranjera, y muy distinto también de la que sentiría al verse frente a un individuo que afirmara sencillamente «querer» un plato de lodo. Si bien las excentricidades particulares que presentan contenidos enteramente opacos pueden tal vez pasarse por alto alguna vez, lo que no puede pasarse por alto son las prácticas comunitarias transmisibles y

con contenidos claros; y la recepción de un deseo transmisible implica un desafío a forjar alguna expresión comunicable de la percepción de la deseabilidad que se halla en el objeto de ese deseo.

2.3 ¿Otra concepción falsa?

Podemos apreciar ahora la idea instructiva mencionada al inicio del capítulo (p. 47) respecto de lo que el deseo es. Esta idea supone que el caso central consistente en que un agente desea algo está constituido por el hecho de que el agente posee una caracterización de deseabilidad del objeto de su deseo, y que hay un caso derivado constituido por el hecho de que posee meramente una percepción de la deseabilidad de ese objeto. Es decir, en un caso de tipo central, el hecho de que un agente desee algo equivale a que cuenta con una caracterización del objeto de deseo, la cual le sirve, a su entender, para llevar a buen término el cuestionamiento en torno a por qué quiere ese objeto: una caracterización del objeto de deseo que le sirve, a su entender, para comunicar la deseabilidad de ese objeto. La inclusión de la frase «a su entender» no indica ninguna relativización en la noción misma de una caracterización de deseabilidad, sino que se emplea únicamente para consignar que se admite la posibilidad de error. En estos casos centrales —según reza la idea— desear consiste en tener una presunta caracterización de deseabilidad *y punto*.

La objeción obvia es que se ha sustituido la concepción clásica falsa y tendenciosa por otra claramente errónea, pues en el discurso ordinario abundan, sin duda, observaciones del tipo «sé que hacer tal o cual cosa es deseable, pero no quiero hacerla». No es más convincente negar esto que la afirmación de que ese mismo discurso está plagado de observaciones de la forma «no quería hacerlo, pero lo hice». Sin embargo, ¿acaso este último fenómeno revela que la ortodoxia vigente sobre las razones para actuar está equivocada? No de una manera directa, pues quien esté de acuerdo con esa ortodoxia intentará justificar la problemática significación del fenómeno mediante una exposición de las diferentes cosas que los hablantes pueden querer decir cuando dicen lo que dicen. Exactamente la misma estrategia está al alcance de quien defienda la concepción del deseo que aquí discutimos.

Comenzaré mencionando algunos de los diagnósticos posibles más obvios de lo que sucede cuando alguien dice: «sé que hacer tal o cual cosa es deseable, pero no quiero hacerla». Con esa aseveración, el hablante puede estar poniendo de manifiesto su adhesión a alguna idea cuasifilosófica sobre los objetos posibles de los deseos humanos, cierta restricción

de esos objetos posibles en términos, por ejemplo, del placer o del interés propio: «Ahí no hay nada para mí; por deseable que sea *no puedo* querer hacer eso.» (Tal restricción se encuentra estrechamente vinculada con una doctrina que desempeña un papel importante en la filosofía moral de Kant.) O una vez más, una aseveración de este tipo puede ser la manifestación de una falsa concepción general del hablante sobre la naturaleza del deseo mismo, una identificación general del deseo con, digamos, un «impulso sentido»: «No siento nada hacia ello; por lo tanto, por deseable que sea, no se da el caso de que quiera hacer eso.» Una posibilidad más compleja consiste en entender las palabras del hablante acerca de lo que es deseable en términos de una mera «gradación» o «clasificación» según cierto esquema de valores que, de hecho, no acepta —sean éstos los valores de la sociedad burguesa, de la Escuela o del Ministerio de Agricultura y Pesca— (véase Urmson 1946). En ese caso, la precisión de la manera de expresarse *requeriría* una relativización explícita de la afirmación de «deseabilidad». Y, por último, están los casos en los cuales lo que está en juego es un contraste general entre las consideraciones morales y las de otra especie: la aseveración a la que se da voz en tales casos equivale a la idea de que los deseos que surgen de consideraciones no morales pesan más que los que surgen de consideraciones morales.

Además de estas posibilidades más obvias, una observación de la forma «veo que hacer tal o cual cosa es deseable, pero no quiero hacerla» puede ser la manifestación de una concepción, que quizá se tenga sólo tácitamente, a la que llamaré concepción del *carácter atomista de la percepción de deseabilidad*: esto es, en términos generales, la concepción de que una vez que se ha observado lo que parece ser un rasgo deseable de, por ejemplo, una acción, la consideración de otros rasgos indeseables de esa acción no es capaz de suprimir nunca la percepción inicial de deseabilidad. Un examen ulterior de la acción podría dar como resultado el reconocimiento de rasgos indeseables de la acción, los que, según se estima, cuentan más que la deseabilidad percibida originalmente; pero la percepción original de deseabilidad subsiste. Según esta concepción, las cosas sólo podrían ser distintas si el reconocimiento de los rasgos indeseables sirviera para cegar al agente ante el aspecto deseable.

¡Qué campo de batalla sería la mente humana si esta concepción atomista fuera correcta! Pues si pensamos atomistamente, parecerá verosímil decir con respecto a una acción cualquiera: «Hay algo que decir en favor y algo que decir en contra de hacer eso —y también en favor y en contra de no hacerlo—» (Davidson 1969: 35). Y toda acción, excepto las impul-

sadas por una luz cegadora, será considerada el resultado de una batalla librada entre deseos en conflicto –una batalla en la que, por lo general, habrá muchas tropas en ambos lados del frente–.

Es instructivo comparar dos clases diferentes de casos que hablan contra esta concepción atomista de la percepción de deseabilidad. Una de ellas se deriva de la idea, conocida desde la época de Aristóteles, de que las fronteras entre cualquier virtud dada y los vicios que la flanquean son vagas, y de que no hay reglas directamente aplicables en la elucidación del contenido de la virtud que sirvan para resolver todos los casos discutibles. La diferencia entre una acción valiente y una temeraria no necesariamente se advierte de inmediato, ni siquiera para una mirada atenta y una mente lúcida. En muchos de estos casos discutibles, el agente puede verse llevado a revisar su juicio inicial de que cierta acción particular sería realmente un caso de, digamos, valentía, al percatarse, por la consideración de otros aspectos de la acción contemplada, de que ejecutarla equivaldría a descuidar las exigencias de, digamos, la justicia. Convencido de que *en este caso* los dictados de la justicia y la valentía no están en conflicto, el agente se ve llevado a retractarse de la caracterización de deseabilidad que expresó en un principio acerca de esa acción: acepta que no sería valiente realizarla. Pero en la otra clase de casos, hay un sentido según el cual el agente no se ve literalmente llevado a retractarse de su «caracterización de deseabilidad» inicial para la acción contemplada; ocurre más bien que, a la luz de la consideración de otros aspectos de la acción contemplada, el agente ya no ve que esa caracterización constituya una caracterización de *deseabilidad* de la acción. Pensando desde un punto de vista atomista, todavía puede aceptar que hay una «caracterización de deseabilidad» de la acción contemplada; al considerar la acción en su totalidad, por así decirlo, no puede ver nada deseable en *ella* (compárese –y contrástese– con McDowell 1978: 26).

Por lo tanto, *con el respeto que me merece* la concepción atomista de la percepción de deseabilidad, parece que uno puede ver que cierto rasgo de una acción es deseable al prestar atención sólo a ese rasgo y, sin embargo, llegar a verla como carente de toda deseabilidad dentro del contexto específico de la acción en su conjunto. Ahora bien, cuando alguien dice algo de la forma «veo que hacer tal o cual cosa es deseable, pero no quiero hacerla», podría suceder que la primera parte de su juicio se derive de la adopción de un punto de vista atomista, mientras que la segunda parte sea resultado de su contemplación de la acción en su conjunto. El hablante se halla en parte bajo la influencia de la concepción atomista: pero sólo en parte, porque de no ser por ello, según la concepción del

deseo que estamos examinando aquí, tendría el deseo apropiado, aunque éste sea un deseo de menor peso.

Sin duda, hay otras posibilidades abiertas para quien quiera eliminar la presunta fuerza de las observaciones de la forma «veo que hacer tal o cual cosa es deseable, pero no quiero hacerla»; pero ya se ha dicho lo suficiente para poner en duda esta *clase* de objeción a la concepción del deseo que estamos discutiendo.

Otra objeción a esa concepción es la siguiente: a grandes rasgos, la identificación de desear con creer deseable pasa por alto la distinción entre deseos y meros anhelos. Quizás poseer una caracterización de deseabilidad cubre por lo menos los casos centrales de anhelo de que suceda tal y cual cosa; pero el deseo exige la aparición de un elemento más, sea de sentimiento, de tendencia o de una combinación de esos elementos. Al menos así se expresa la objeción.

No se puede negar que en nuestro idioma existen las formas verbales, claramente distintas, «desear» y «anhelar».[10] Pero de ello no se sigue ni que haya dos clases diferentes de estados mentales, deseos y anhelos, ni que desear equivalga a anhelar *más* otra cosa. En términos muy generales —en el nivel de la «semántica popular»— la situación parece ser ésta: el verbo «anhelar» [*to wish*] se utiliza como expresión de deseo en ciertas circunstancias específicas, entre las cuales la más central es aquella en que el agente cree que la realización del objeto de su deseo se halla fuera de su alcance (pensemos, por ejemplo, en deseos dirigidos al pasado.) Éste sólo es el caso básico: pues el fenómeno que hay que explicar es lioso. Pero quien quiera emplear ese lío como una objeción a la concepción del desear que estamos contemplando, también tiene que dar algunas explicaciones: quizá, muy en particular, tiene que explicar el hecho de que las mismas consideraciones que debilitan en un contexto dado la pretensión de sinceridad de un hablante cuando expresa sus «deseos», también sirven dentro del mismo contexto para debilitar la pretensión de sinceridad del hablante cuando expresa sus «anhelos». Da lo mismo si digo «anhelo que esa ventana esté abierta» o «quiero que esa ventana esté abierta»: si alguien se ofrece a abrirla y mi reacción es rechazar el ofrecimiento, lo que va a pasar es lo mismo.

Esto es en el nivel de la semántica popular; las cosas no son muy diferentes cuando se permite que la teoría entre en escena. Puede ser que para algunos propósitos teóricos sea útil una distinción nítida entre deseos y

10. En inglés: *to desire* y *to wish*, respectivamente. [N. de los tt.]

anhelos. Consideremos por tanto uno de esos casos. En sus conferencias «William James», Richard Wollheim expresó lo siguiente:

> Cuando deseo algo, lo anhelo y no meramente lo deseo; y porque lo deseo, tiendo a imaginar (en el modo apropiado) mi deseo cumplido; y cuando lo imagino cumplido, para mí es como si ese deseo estuviera cumplido. (Wollheim 1984: 90)

A pesar de las subsiguientes precauciones de Wollheim sobre la cuestión, la dependencia conceptual de esta noción de anhelar respecto de la de desear parece clara: la objeción que considera desear como anhelar *más algo* coloca conceptualmente el carruaje delante del caballo.

2.4 Objetos del deseo profundamente oscuros

Anteriormente nos hemos visto inducidos a reconocer la existencia de un punto de transición en la búsqueda de comprensión por parte de un investigador, una transición en la cual éste llega a ver que una actividad es «de algún modo deseable» y a actuar en consecuencia, aun cuando todavía no haya alcanzado una caracterización de deseabilidad para la actividad en cuestión. Ese punto de transición bloqueó toda identificación *simpliciter* del desear con el poseer una caracterización de deseabilidad. Pero ahora se plantea la objeción de que ese hablar de ver que una actividad es «de algún modo deseable» no equivale más que a registrar el hecho de que la persona actúa en consecuencia; o sea, ¡que tiene el deseo pertinente! La concepción del deseo que examinamos amenaza con derrumbarse en la vacuidad. Y siendo esto así, podría proseguir la objeción, no tenemos ahora ninguna razón en absoluto para poner en duda la categoría de deseos de los «deseos» que los dioses indujeron en el Sísifo de Richard Taylor mediante la implantación de una sustancia en sus venas (capítulo 1, pp. 42–46). Además, aquí se ha aceptado que tiene un deseo la persona que realiza, por ejemplo, una actividad ritual debido a una percepción de deseabilidad, pero que carece todavía de una caracterización de deseabilidad para esa actividad (p. 53); pero ese deseo «infeccioso» parece hallarse tan aislado de la interacción con los otros deseos y creencias del sujeto como lo está el supuesto deseo que Taylor atribuye a Sísifo.

Estas inquietudes no se pueden descartar sin más ni más haciendo referencia a nuestro conocimiento de cómo se indujo el supuesto deseo de Sísifo. Más bien habría que subrayar los siguientes detalles. No tenemos razón alguna para creer que un investigador forastero que se uniera

a Sísifo en su actividad llegaría a compartir, independientemente de su proyecto de investigación, el supuesto deseo de Sísifo. No tenemos razón alguna para creer que exista una percepción de deseabilidad acompañada de una caracterización oculta de deseabilidad que se relaciona con otras acciones practicadas por Sísifo. Y no tenemos razón alguna para creer que haya una percepción de deseabilidad acompañada de una oculta caracterización de deseabilidad que, si fuera descubierta por Sísifo, podría conducirlo, dependiendo de sus otros deseos y creencias, a alguna entre una gama ilimitada de acciones o a alguna forma de deliberación protopráctica. Piénsese cuán diferente sería el caso si el efecto de la sustancia implantada fuera fomentar el desarticulado interés de Sísifo por su desarrollo físico. Su deseo resultaría entonces transmisible (a algunas personas, por lo menos); probablemente, se trasluciría en otras acciones suyas, y aunque no fuera así, parecería probable que lo hiciera si Sísifo llegara a reconocer que eso fue en realidad lo que lo había movido a dedicarse a empujar rocas. Más aún, si llegara a tal reconocimiento, estaría entonces en posición de emprender alguna deliberación protopráctica que pudiera desembocar en una apreciación modificada del lugar de ese deseo dentro de su visión del mundo, una apreciación modificada que podría tener innumerables ramificaciones en su conducta. Lo mismo ocurriría con nuestro investigador de la cultura ajena mientras se hallara en la etapa de transición. Ha llegado a ver que la actividad ritual es «de algún modo deseable»; se le ha transmitido el deseo aunque no cuenta con ninguna caracterización de deseabilidad de la actividad ritual. No obstante, es probable que la novedosa percepción de deseabilidad se manifieste en formas nuevas o modificadas de su comportamiento. Y aun cuando no se manifestara, es probable que lo hiciera si el investigador llegara a reconocer la caracterización oculta de deseabilidad para esa actividad. Y tal reconocimiento admitiría una deliberación protopráctica encaminada a la apreciación del lugar que ese deseo merece dentro de su concepción del mundo. La idea no es aquí únicamente que la caracterización que hace Taylor del caso de Sísifo calle todas estas otras consideraciones que podrían comenzar a fundamentar el empleo del concepto de deseo para ese caso; más bien, lo importante para nuestros propósitos actuales es que la apreciación de esas otras consideraciones libera de la apariencia de vacuidad a la concepción del deseo que estamos considerando.

No obstante, creo que esta concepción del deseo afronta una dificultad más formidable. Las percepciones de deseabilidad se introdujeron para dar razón de la existencia del punto de transición: de casos en que el

agente tiene un deseo pero carece de una caracterización de deseabilidad adecuada para el objeto (pp. 53–54). Una ilustración sencilla de esos casos sería una situación en que una necesidad o carencia no reconocida diera directamente origen a un deseo. En muchos de esos casos, el agente no necesita reflexionar mucho para reconocer la necesidad o la carencia que surge y llegar así a una caracterización de deseabilidad adecuada. Pero se necesite o no mucha reflexión al respecto, la discusión de la última objeción ha puesto en claro que el contenido del hablar sobre las percepciones de deseabilidad termina por apoyarse en su conexión con el hablar sobre las caracterizaciones de deseabilidad *disponibles*. Aun así, en un grupo importante de casos está lejos de resultar claro que tal relación exista: pues en tales casos encontramos objetos de deseo *profundamente* oscuros.

Considérese lo que podría denominarse *actividades de sustitución*. Éstas abarcan desde el caso del marido solitario que, lejos de su hogar, besa la foto de su esposa antes de dormirse, pasando por aquel en que, ya solos, proseguimos la violenta disputa que acaba de hacernos abandonar la compañía de otro, hasta por lo menos el caso de quien se lava las manos compulsivamente. Russell dio un ejemplo característico en *The Analysis of Mind*:

> Supongamos que nos han dejado plantados de tal modo que nuestra vanidad ha quedado herida. Nuestro deseo impulsivo natural se asemejará al que se expresa en el poema de Donne:
>
> *Cuando por tu desprecio, oh Asesina, he muerto*
>
> en el cual explica cómo, siendo él un fantasma, perseguirá a la pobre dama, impidiéndole tener un momento de reposo. Pero hay dos cosas que se oponen a que nos expresemos de una manera tan natural: por un lado, nuestra vanidad, que no reconocerá cuán profunda es la herida; por el otro, la convicción de que somos personas civilizadas y humanas que no podrían complacerse en un deseo tan ordinario como la venganza. En consecuencia, experimentaremos un desasosiego que, al principio, parecerá no tener ningún objetivo, pero que finalmente se resolverá en un deseo consciente de cambiar de profesión, de dar la vuelta al mundo o de ocultar la identidad propia e irnos a vivir a Putney [... Sin embargo], el viaje nos parecerá decepcionante y el Oriente menos fascinante de lo que habíamos esperado —a menos que un día nos enteremos de que a la malvada también la han dejado plantada. (Russell 1921: 73)

Un ejemplo quizá más exótico se halla en *Nuestra señora de las flores* de Genet:

> Divina le presentó Nuestra Señora a ella. Unos días más tarde, le mostró, niña decente como era, una pequeña fotografía «fotomática» del asesino.
>
> Mimosa toma la foto, se la pone en la punta de la lengua y se la traga.
>
> «Adoro a esa Santa Señora vuestra. Estoy comulgándomela.» (Genet, 1963: 204)

Los ejemplos no son tan cotidianos, pero el fenómeno general sí lo es. En pocos de estos casos resulta fácil descubrir una caracterización de deseabilidad convincente para la actividad; e incluso cuando parece fácil, es correspondientemente difícil hallarle mucho sentido a la creencia que se le atribuye en consecuencia al agente en la racionalización de su acción.

Es importante ser claros en lo tocante a las dificultades distintivas que plantean tales casos. En algunos otros casos perfectamente comunes de «error activo» por parte de un agente, éste posee una caracterización de deseabilidad directa y adecuada para cierto tipo de acción, y al mismo tiempo tiene una creencia comprensiblemente equivocada acerca del carácter de la acción que ejecuta; en otros casos, junto con la creencia correcta acerca del carácter de la acción que ejecuta, tiene también una presunta caracterización de deseabilidad inteligiblemente equivocada para el tipo de acción ejemplificado. En ambas clases de casos, el agente tiene cierta caracterización de deseabilidad inteligible que se pone de manifiesto en una amplia variedad de comportamientos en diferentes circunstancias. E incluso si existe un objeto del deseo temporalmente oscuro —incluso si de momento el agente solamente tiene una percepción de deseabilidad—, hemos visto cómo esa percepción, en virtud de sus vínculos con alguna caracterización de deseabilidad disponible, puede ponerse de manifiesto en una amplia gama de comportamientos en diferentes circunstancias (pp. 58–59). Pero las actividades de sustitución no ejemplifican en general ninguna de estas clases de casos.

La dificultad que hay aquí puede plantearse en forma de dilema: o vinculamos la noción de percepción de deseabilidad con la de una caracterización de deseabilidad disponible, en cuyo caso por lo menos algunas actividades de sustitución representan casos de deseo sin percepción de deseabilidad; o bien la vinculamos con la de deseo, despojándola, *en estos casos*, de los nexos que tiene con la noción de una caracterización de deseabilidad disponible, y de este modo utilizamos la noción de percepción de deseabilidad meramente para registrar la presencia del deseo. Luego, si tomamos la concepción del deseo que estamos considerando como una tesis ilimitadamente general sobre la naturaleza del deseo, esa concepción es o falsa o vacía.

Prefiero emplear la expresión «percepción de deseabilidad» de acuerdo con el primer cuerno de este dilema, por lo cual me veo llevado a rechazar por falsa la tesis que consideramos sobre lo que el deseo es. Pero esto no implica negar la importancia de las conexiones entre la noción de deseo, por un lado, y las nociones de caracterización de deseabilidad y percepción de deseabilidad, por el otro. En consecuencia, tampoco hay nada de lo dicho que niegue la importancia del progreso logrado en lo tocante a ampliar el círculo de conceptos dentro del cual el concepto de deseo ocupa su lugar.

2.5 El deseo mismo

En el primer capítulo se presentaron razones para apoyar el rechazo de la concepción clásica del deseo. Hasta aquí, en el presente capítulo se han aducido razones contra una alternativa a esa concepción sobre lo que es el deseo (y más adelante surgirán otras razones contra esa concepción alternativa). Pero, ¿qué es, entonces, el deseo?

En *Sendas del significado* expuse las siguientes observaciones que ahora pienso que pueden servir como glosa siquiera para la distinción entre deseo, por un lado, y las caracterizaciones y percepciones de deseabilidad (según se han entendido hasta el momento aquí), por el otro:

> En su trabajo sobre la intención, la profesora Anscombe ha trazado una distinción amplia entre dos *tipos* de estado mental; la creencia fáctica es el principal ejemplar de un tipo, y el deseo es el ejemplar principal del otro tipo (Anscombe, *Intention*, 2). La distinción se da en términos de la *dirección de adecuación* [*direction of fit*] de los estados mentales con el mundo. Las creencias buscan la verdad y que sean verdaderas equivale a que se ajustan al mundo; la falsedad es un defecto determinante en una creencia y las creencias falsas tendrían que ser descartadas; las creencias tendrían que ser cambiadas para que se ajusten al mundo y no al revés. Los deseos apuntan a la realización y su realización ocurre cuando el mundo se ajusta a ellos; el hecho de que el contenido indicativo de un deseo no sea realizado en el mundo no es todavía un defecto *del deseo*, ni tampoco una razón para descartarlo; el mundo, *grosso modo*, tendría que ser cambiado para ajustarse a nuestros deseos y no al revés. (Platts 1979: 256-257 [1992: 369-370])*

Ésta es una glosa metafórica, y me parece dudosa la probabilidad de una captación más literal de la distinción así trazada. (No me sorprende que

* Las referencias entre corchetes remiten a las versiones en castellano de las obras citadas por el autor y de las cuales tomamos, a veces con modificaciones, la traducción. [N. de los tt.]

un intento reciente por captar esa distinción esté basado en una falsa concepción disposicional del deseo (Smith 1987).) Pero el carácter metafórico de esta forma de marcar la distinción puede amenazar con volver problemática cualquier tesis subsecuente sobre la ejemplificación efectiva de la distinción. Por esta razón, he preferido la ruta más larga de establecer una distinción real por medio de la consideración de las realidades de las actividades de sustitución. Habiendo establecido así la distinción real, no veo mayores problemas en glosarla luego en términos metafóricos.

Los resultados positivos obtenidos hasta ahora en esta indagación sobre la naturaleza del deseo pueden parecer un tanto cuanto magros. Pero, ¿por qué habría de suponerse que en términos positivos haya mucho que decir, *en este nivel de abstracción*, acerca de un concepto tan general y tan ampliamente usado?

Lo que ahora propongo hacer es trabajar con miras a una taxonomía filosófica de los deseos, una descripción filosófica de las diferentes *clases* generales de deseos. Como cualquier descripción de esa índole, ésta tendrá que ser a la vez detallada y verídica; el interés filosófico más general de los detalles descriptivos surgirá después en el capítulo siguiente.

2.6 Algunas distinciones iniciales

En *The Possibility of Altruism* (1970: 29–32), Thomas Nagel distinguió los deseos motivados de los no motivados. El peso de su explicación de esta distinción exhaustiva recayó en su elucidación de los deseos motivados. En realidad, parecía haber tres elucidaciones diferentes intentadas. La tesis inicial era que a los deseos motivados «se llega por decisión y después de una deliberación», mientras que la tesis final parecía ser la de que los deseos motivados se atribuyen de la misma manera derivada que las creencias en los principios elementales de la lógica. Pero entre ambas había una tercera tesis que quizás eludía los errores de la primera, el carácter evasivo de la última y la marcada apariencia de una tensión entre estas dos:

> [S]i mi deseo es motivado, su explicación será la misma que la explicación de [la búsqueda de la meta por parte del agente], y no es de ninguna manera obvio que deba intervenir un deseo en esta última explicación. Aunque sin duda se admitirá en general que algunos deseos están motivados, la cuestión es si siempre se encuentra otro deseo detrás del motivado, o si en ocasiones la motivación del deseo inicial no implica ninguna referencia a algún otro deseo no motivado. (*Ibid.*)

Con respecto a esta cuestión, la postura de Nagel es clara:

> Si aplicamos estas observaciones a la cuestión de si los deseos figuran siempre entre las condiciones necesarias de las *razones* para actuar, resulta obvio que no hay razones para creer que así sea. A menudo, los deseos que un agente experimenta necesariamente al actuar tendrán las mismas motivaciones que la acción. Si el acto está motivado por razones emanadas de ciertos factores externos, y el deseo de ejecutarlo está motivado por esos mismos factores, es obvio que el deseo no puede contarse entre las condiciones para la presencia de esas razones. (*Ibid.*)

Así, dentro de la clase de los «deseos motivados», se distingue primero una subclase más o menos en los siguientes términos: para cualquier deseo que esté dentro de esta subclase, una explicación completa de la presunta deseabilidad de su objeto hará referencia, lógicamente, no a la existencia de ese deseo mismo, sino más bien a la existencia de algún *otro* deseo que el agente tenga. Además, en cualquier caso particular de este tipo, el agente puede ser capaz de defender la tesis de que la explicación en cuestión sirve para captar la deseabilidad del objeto haciendo referencia tanto a la existencia de su otro deseo como a los hechos pertinentes sobre el papel que el objeto desempeña potencialmente en el proyecto de satisfacer ese otro deseo. Dos ejemplos que parecen obvios de deseos de esta clase son, primero, aquellos en que el objeto del deseo es un *medio* para conseguir el otro objeto deseado, y segundo, aquellos en que el objeto del deseo es *parte constitutiva* del otro objeto deseado. Pero más problemática y más importante es la afirmación de que hay otra subclase de deseos motivados de los que debe darse razón en términos que, intentando evitar referirse a otros deseos, aludan sólo a «razones emanadas de ciertos factores externos».

El interés central de Nagel era en efecto la posibilidad del altruismo; sin embargo, una parte estratégicamente crucial de su argumento encaminado a establecer esa posibilidad estaba dedicada a demostrar que por lo menos algunos deseos prudenciales son miembros de la subclase problemática de los deseos motivados que acabamos de mencionar. A este respecto, Nagel estaba preocupado por mostrar la manera en que un agente podría sentirse racionalmente motivado por la idea de que en el futuro tendrá una razón para propiciar cierto estado de cosas, independientemente de cualquier deseo actual no motivado. Esa posibilidad puede ser explicada, sostenía Nagel, haciendo referencia a una concepción que tenemos del tiempo, a saber, que todos los tiempos —pasado, presente y

futuro— son igualmente reales. En los términos más sumarios posibles: para quien adopte cabalmente esa concepción, su juicio de que tendrá una razón para propiciar cierto estado de cosas lo compromete, independientemente de cualquier deseo actual no motivado, con el juicio de que ahora tiene una razón para propiciar dicho estado de cosas; y la aceptación de este último juicio equivale, en igualdad de circunstancias, a estar motivado a propiciar ese estado de cosas. No obstante, cuando una motivación surge de ese modo, para Nagel sigue siendo cierto que puede atribuirse al agente un deseo presente correspondiente: que el agente tenga el deseo correspondiente simplemente se sigue del hecho de que las consideraciones lo motivaron. El deseo atribuido así es, entonces, un miembro de la subclase problemática de los deseos motivados.

Nuestro interés actual no se dirige a los detalles de la explicación de Nagel sobre la prudencia, pero no está de más una advertencia sobre la aceptación aparentemente difundida que ha tenido la misma. No está muy claro hasta qué grado otros autores posteriores han adoptado detalles específicos de la explicación de Nagel sobre la motivación prudencial. Philippa Foot, por ejemplo, aparentemente se adhirió a ésta cuando sostuvo que las razones prudenciales ofrecen los contraejemplos más obvios de la tesis según la cual todas las razones para la acción dependen de los deseos del agente. El deseo correctamente atribuido a un agente que cuenta con una motivación prudencial, no puede, según Foot, ser la base de las razones para actuar; por el contrario,

> [l]o que aquí tenemos es un uso de la palabra «deseo» que indica una dirección motivacional y nada más. Se puede comparar con el uso de «quiero» en «quiero ϕ», donde sólo se implica la intencionalidad. ¿Puede el *querer*, en este sentido, crear la razón para actuar? Parece que no. Pues en primer lugar, los deseos de los que estamos hablando ahora deben atribuirse al agente sólo en caso de que éste se vea movido a la acción, o de que se viera movido a ella en ausencia de razones o causas que se le opusieran, y constituye un error fijar límites correspondientes a la esfera de las razones para actuar. Además, se da una falsa explicación hasta de los casos en que se produce la acción. Pues lo que entonces sucede es que un hombre se ve movido a la acción por el reconocimiento de que tiene una razón para actuar. Esto sería imposible si no hubiera razones que reconocer hasta que el agente haya sido motivado. (Foot 1972b: 149–150 [1994: 174–175])

Sin embargo, en una nota posterior al artículo en que parece adoptar la explicación de Nagel, Foot expresa cierta inclinación por el punto de vista de que todas las razones para actuar dependen o bien de lo que

favorece el interés del agente, o bien de sus deseos (*ibid.*, p. 156). Pero ello indica que su distancia de Nagel es realmente grande, puesto que, según la explicación de éste, las razones prudenciales son razones que se relacionan con todo tipo de previsión para el futuro y no tienen, en general, ninguna conexión particular con los intereses futuros del agente.

Un escritor más reciente, que parece aproximarse más a Nagel en su elucidación de la distinción entre deseos motivados y no motivados, y de la distinción dentro de la clase de los deseos motivados, es John McDowell:

> La especificación completa de una razón [...] debe contener los elementos suficientes para revelar la luz favorable bajo la cual el agente contempló la acción proyectada. Tendemos a suponer que, en líneas muy generales, tal es el efecto de la inclusión de un deseo [...] Sin embargo, parece ser falso que el poder motivador de todas las razones se derive de que incluyan deseos. Supongamos, por ejemplo, que explicamos la ejecución de cierta acción por parte de una persona adjudicándole a ésta conciencia de algún hecho que vuelve probable (desde su punto de vista) que actuar de esa forma favorecerá su propio interés. Hacer referencia a su visión de los hechos puede bastar, por sí solo, para mostrarnos la luz favorable bajo la cual se le presenta su acción. Sin duda, le adjudicamos un deseo apropiado, acaso con miras a su propia felicidad futura. Pero el compromiso de atribuirle ese deseo es simplemente una consecuencia de que hayamos supuesto que actúa como lo hace por la razón que citamos [...] (McDowell 1978: 14-15)

Y en un escrito posterior señala:

> Las explicaciones de las acciones en términos de razones funciona mediante la revelación de la luz favorable bajo la cual el agente contempló lo que hizo (o por lo menos lo que intentó). En [«Are moral requirements hypothetical imperatives?» («¿Son los requisitos morales imperativos hipotéticos?»)], distinguí entre los casos en los cuales esto sólo se logra aludiendo a un deseo o dándolo por sentado, y los casos en los cuales la mención de un estado cognoscitivo (la concepción que tiene el agente sobre la situación) basta para mostrar la luz favorable bajo la cual el agente contempló su acción. (McDowell 1982: 301)

Por lo tanto, para McDowell, como para Nagel, la atribución de un miembro de la subclase problemática de los deseos motivados es simplemente una consecuencia de que consideremos que el agente actúa como lo hace por la razón que citamos en la explicación de su acción; pero el peso de la explicación —de la explicación tanto de la acción como del deseo así

atribuido— recae, según McDowell, en la «visión de los hechos» del agente, en «un estado cognoscitivo (la concepción que tiene el agente sobre la situación)». De este modo, el discurso de Nagel acerca de las «razones emanadas de ciertos factores externos» se interpreta en términos de los «estados cognoscitivos», pensando quizá que estos últimos «emanan» de los hechos en cuestión (los *factores* externos).

2.7 Algunas cuestiones iniciales

Surgen de inmediato dos enigmas. El más burdo es el siguiente. Tanto para Nagel como para McDowell (y también para Foot) el deseo nunca queda completamente fuera del cuadro de la agencia: aun cuando el peso de la explicación o de la motivación recaiga, desde el punto de vista de McDowell, en algún estado cognoscitivo, en la «visión de los hechos» del agente, ha de atribuírsele, sin embargo, un deseo al agente. El deseo así atribuido es un deseo motivado «apropiado», y se considera que su atribución «es simplemente una consecuencia de que hayamos supuesto que la persona actúa como lo hace por la razón que citamos». El enigma aquí resulta obvio. Los intentos más conocidos de fundamentar esa parte de la ortodoxia contemporánea que afirma que una especificación completa de la razón para actuar que tiene un agente debe hacer referencia a sus deseos (o «actitudes favorables») así como a sus creencias, se efectúan mediante la consideración de ejemplos en los cuales se supone que las creencias del agente son claramente neutrales respecto de la motivación —característica que a menudo queda asegurada por medio de un alto nivel de abstracción, a través del silencio, en relación con el agente y su situación—. Si los ejemplos se montan de esta manera, entonces parecerá claro que lo que el agente se sienta impulsado a hacer, a intentar hacer o a considerar seriamente hacer, dependerá de la naturaleza de sus deseos todavía no especificados. La especificación de los contenidos de esos deseos parecerá ser *el* medio de revelar el bien imaginado a cuya luz el agente actúa o piensa actuar. Pero tal resultado sólo se asegura suponiendo que las creencias atribuidas inicialmente al agente son neutrales en lo que toca a cuestiones de motivación, y esto es precisamente lo que se supone que no ocurre en los casos destacados por Nagel y McDowell dentro de su tentativa de establecer la existencia de un modelo distintivo de explicación de la acción. En los casos preferidos por Nagel y McDowell, el peso de la explicación debe recaer en cierto «estado cognoscitivo» del agente: ¿por qué entonces habrían de persistir en la atribución de deseos en esos casos?

Hay en verdad una distancia considerable entre los puntos de vista de Nagel y McDowell, por un lado, y la plena ortodoxia sobre las razones para actuar que ellos atacan, por el otro. Esa ortodoxia no se limita a afirmar que las razones para actuar siempre incluyen deseos además de creencias; también sostiene que los deseos son, en todos los casos, los que llevan el peso motivacional-*cum*-explicativo. Es decir, para esta ortodoxia, los poderes activos implicados en una acción ejecutada por una razón son siempre deseos; son siempre deseos los que motivan y mueven al agente a actuar. El rechazo de este segundo elemento de la ortodoxia marca así una distancia considerable. No obstante, el enigma subsiste: ¿por qué, exactamente, debe mantenerse siempre al deseo dentro del cuadro?

El otro enigma es menos obvio. Al menos en un nivel intuitivo, parece que podemos distinguir, para un agente dado en una situación específica, tres diferentes *usos* de los enunciados relativos a su situación. (No todos los usos pueden distinguirse por referencia al tipo de vocabulario que aparece en los enunciados usados; acaso ninguno.)

(1) La atribución que el agente hace a sí mismo de un deseo de realizar la acción ya de hecho ejecutada bajo cierta descripción específica de esa acción.
(2) El ofrecimiento, por parte del agente, de una caracterización de deseabilidad de la acción realizada, caracterización que él supone adecuada para comunicar la deseabilidad que a sus ojos posee la acción.
(3) El ofrecimiento, por parte del agente, de una caracterización objetiva de la acción realizada, caracterización que, sin ser en sí misma una caracterización de deseabilidad, podría servir, sin embargo, como una razón en apoyo de alguna caracterización de deseabilidad.

Ahora bien, ¿cuál, si alguno, de estos usos interviene en la especificación apropiada de la «concepción de la situación» que tiene el agente, de su «estado cognoscitivo» (motivador)? ¿Cuál, si alguno, interviene en una especificación apropiada de su «situación», de los «hechos»? ¿Y qué explicación puede darse, si puede darse alguna, de la idea de que lo primero «emana» de lo segundo; idea que resulta esencial para el pensamiento de que lo primero es verdaderamente un estado *cognoscitivo*?

2.8 Refinamiento de las distinciones iniciales

El primer enigma se relaciona, pues, con la persistencia de Nagel y McDowell en atribuirle un deseo al agente aun en casos en los cuales el peso

explicativo o motivacional recae en la referencia a cierto estado cognoscitivo, a la «concepción de los hechos» del agente. Sospecho que las ideas determinantes implicadas se desarrollaron como sigue. En primer lugar, hubo una aceptación de la explicación metafórica del deseo en términos de su distintiva «dirección de adecuación» con el mundo. Luego, dentro del alcance de esa aceptación, surgió otra idea según la cual aun cuando el peso explicativo o motivacional recaiga en un estado mental con una «dirección de adecuación» con el mundo de tipo creencia, la acción pone de manifiesto, sin embargo, la presencia de un estado de deseo con *su* distintiva «dirección de adecuación» con el mundo. Esquemáticamente, la argumentación para defender esta otra idea podría haber sido ésta: *cualquier acción (excepto las omisiones intencionales) es, y debe suponerse que es, un cambio en el mundo; pero cualquier cambio intencional semejante exige que se dé un estado mental con una «dirección de adecuación» del tipo que distingue al deseo.* Supongo que ésta es también la línea de argumentación que hace que Foot pueda hablar de una «dirección motivacional» (p. 65); supongo también que lo que más probablemente oculta esa línea de argumentación es la adhesión a una concepción del deseo como poder activo —una concepción del deseo como un fenómeno atomista cuasimecánico, como un empujón interior carente de contenido proposicional—. Pero debe quedar claro en todo caso que la línea de argumentación en cuestión difícilmente podría diferir más de los argumentos más comunes encaminados a establecer el mismo resultado que fueron mencionados anteriormente (p. 67); también creo que ha quedado claro que esta línea de argumentación se opone a la concepción del deseo que fue anteriormente rechazada sobre la base más particular de la realidad de las actividades de sustitución (pp. 61–62).

El otro enigma se refiere a tres usos posibles diferentes que el agente puede hacer de los enunciados: (1) atribuirse a sí mismo el deseo de realizar una acción ya ejecutada de hecho; (2) ofrecer una caracterización de deseabilidad de la acción, y (3) ofrecer una caracterización de la acción que sirva de apoyo a alguna otra caracterización de deseabilidad de dicha acción.

Según la concepción del deseo en que nos estamos basando aquí, glosaremos de nuevo la distinción general entre los usos mencionados en (1) y (2) en cuanto a sus distintivas «direcciones de adecuación» con el mundo. Pero con mayores detalles, una diferencia que a menudo puede producirse entre los usos (1) y (2), que en todo caso son distintos, queda ejemplificada en los casos en que la *especificación* de la acción implicada,

dentro del contenido del deseo que uno se atribuye a sí mismo, es diferente de la que constituye la caracterización de deseabilidad dada para esa acción. Esto puede deberse, simplemente, a que la caracterización de deseabilidad entraña una mayor enumeración de detalles, dirigida a cierto público; o, lo que es más interesante, puede en vez de ello reflejar el acaecimiento de un caso en el cual el agente se encuentra con un deseo para cuyo objeto carece al principio de una caracterización de deseabilidad adecuada, de suerte que busca y supuestamente encuentra tal caracterización.

En lo que respecta al uso (3), pienso que lo «lioso» del fenómeno pragmático impone que todo lo que pueda ofrecerse sea una lista, quizá inevitablemente incompleta, de las clases de casos en que el uso (3) se produce *en cuanto caracterizado de manera distintiva*. (Este énfasis es esencial: muchos casos que podrían confundirse con ese uso no equivalen más que a un grado más nítido de especificación de la caracterización de deseabilidad pertinente.) En todos esos tipos de casos, lo que se dice al ofrecer la razón, así como la manera exacta en que se dice, dependerán de la naturaleza del público particular al que esa razón esté dirigida; y el asunto será diferente en los casos en que los dos «agentes» coexistan en una y la misma persona. No obstante, en términos generales, sospecho que el más común de tales casos sería aproximadamente el siguiente: dos agentes aceptan las mismas caracterizaciones de deseabilidad pero difieren en cuanto a la cuestión de si una caracterización dada se aplica al caso de que se trata debido a cierta diferencia en torno a una cuestión de hecho. Quizá uno de ellos crea que el beneficiario de la acción es un amigo verdadero y leal, mientras que el otro cree que no es más que un compañero «en las buenas». Diversas consideraciones pueden hallarse tras el desacuerdo en torno al hecho; pero sólo dentro de una reducida subclase de tales casos dicho desacuerdo estará fundado en el fenómeno, muy enfatizado, en general, por McDowell, de la existencia de una diferencia (en este caso) entre las *concepciones* de los agentes sobre lo que es la amistad. (Véanse, por ejemplo, sus observaciones acerca de distintas concepciones sobre lo que es ser «tímido y sensible» (McDowell 1978: 21-22).) No todas las divergencias en torno a las creencias expresadas en términos de amistad apuntan a una diferencia de concepción. Sin embargo, es más importante el hecho de que esa subclase de casos se distingue de los casos en los que concepciones divergentes acerca de lo que es la amistad, por ejemplo, dan por resultado diferencias respecto a qué caracterizaciones *generales* constituyen verdaderamente caracterizaciones de deseabilidad; y también se distingue mucho de los casos en que esas concepciones di-

vergentes dan por resultado diferencias con respecto al *peso* asignado a las caracterizaciones de deseabilidad ya acordadas. Esto es: los posibles papeles desempeñados por alguna «diferencia de concepción» constituyen un grupo notablemente heterogéneo. Y obsérvese, como precaución final, que ninguna de las posibilidades mencionadas necesita entrar en juego en el caso de un «escueto» desacuerdo en torno a si, supongamos, una acción dada es ejemplo de generosidad o de extravagancia: las razones pueden agotarse en tales casos.

De suerte que, en la que creo que es la clase más común de casos en los que ocurre un uso distintivo de tipo (3) para el ofrecimiento de razones, la razón aducida se encamina a mostrar la aplicabilidad de una caracterización de deseabilidad acordada a una acción particular, resolviendo alguna diferencia en torno a una cuestión de hecho. El expositor de la razón puede considerar que la diferencia ha surgido sólo porque el otro descuidó o malinterpretó un hecho pertinente. Pero que aduzca esa razón no necesariamente implica que su aplicación original de la caracterización de deseabilidad al caso de que se trate sea resultado de alguna inferencia.

¿Cuál de los usos que se han distinguido interviene en la especificación hecha por el agente del contenido de su «concepción de la situación» —si es que alguno interviene—? ¿Cuál interviene en una especificación apropiada de los «factores externos», los «hechos» —si es que alguno lo hace—? ¿Y qué explicación puede darse —si alguna puede darse— de la idea de que lo primero se «deriva» de lo último, de tal suerte que lo primero se revela como un estado cognoscitivo? Creo que las respuestas que deben darse están claras, al menos en gran parte. En la clase distintiva de casos que estamos considerando, dentro de la subclase problemática de los deseos motivados, el peso motivacional recae en algún estado cognoscitivo del agente, cuya especificación plena basta para revelar el bien imaginado a cuya luz actuó. En los casos más claros, la especificación del agente contiene, por tanto, una caracterización de deseabilidad de la acción de hecho ejecutada. Así pues, los casos en que el agente recurre a un uso del tipo (3), en cuanto caracterizados de manera distintiva, no pueden ser casos en que haya un uso adecuado para la plena especificación de su contenido. Más bien, la especificación que el agente hace de ello requiere de su parte un uso del tipo (2): su especificación del contenido de su estado cognoscitivo se logra mediante la presentación de una caracterización de deseabilidad. Resulta igualmente claro que el pensamiento de que lo que se ha especificado es un estado cognoscitivo exige que pueda darse,

usando la misma caracterización de deseabilidad, una especificación adecuada de los «hechos» de los que se deriva ese estado; es decir, exige que se piense en esa caracterización como un señalamiento de características de deseabilidad «externas» de la acción en cuestión. Dentro de la subclase de los deseos motivados, pues, el agente considera la deseabilidad de la acción en cuestión como *independiente* de su reconocimiento actual de esa deseabilidad.

¿Qué puede decirse acerca de la relación entre el estado de deseo enunciado utilizando el tipo (1) y el hecho de que el agente tenga una caracterización de deseabilidad como la enunciada mediante un uso de tipo (2)? Sin salirnos de la problemática subclase de los deseos motivados, me parece que la mejor propuesta es ésta: en los casos en que se verifica una acción intencional adecuada, el hecho de que el agente haya tenido una caracterización de deseabilidad *provocó, dio lugar a, desembocó razonablemente en* que el deseo se haya hecho manifiesto en esa acción (véase Platts 1980b: 76-78). La relación que toscamente se está señalando allí será un tema recurrente en gran parte de lo que sigue; sin embargo, algunas observaciones negativas iniciales pueden evitar un inminente malentendido. A riesgo de confundir relaciones enteramente distintas, consideremos primero la relación entre una pareja apropiada de creencia-deseo —al margen del tipo de deseo implicado— y la correspondiente intención de actuar. Si en un caso dado lo primero da lugar a lo segundo, la pareja creencia-deseo racionaliza —vuelve por lo menos mínimamente razonable— la intención así formada; pero sería un error muy grave concluir de esto que la intención así formada se formó intencionalmente. De modo similar, sostener que el hecho de que el agente tenga una caracterización de deseabilidad puede dar lugar a un deseo correspondiente, no equivale a sostener que el agente forma intencionalmente el deseo. Consideremos, en segundo lugar, la relación entre una pareja apropiada de creencia-deseo —al margen del tipo de deseo implicado— y la acción correspondiente. Si la acción se verifica, la pareja creencia-deseo racionaliza la acción así ejecutada. Pero hemos visto en el primer capítulo que no es preciso que tal explicación racionalizante se preste a un proyecto predictivo: si en un caso dado la acción no se verifica, no necesita haber ninguna explicación de ese «fracaso», formulada en cuanto a la vida mental y personal del agente. De manera semejante, entonces, afirmar que el hecho de que un agente tenga una caracterización de deseabilidad puede dar lugar a un deseo correspondiente no equivale a afirmar, por ejemplo, que cualquier diferencia

entre los deseos correspondientes *tiene que* reflejar alguna diferencia entre las caracterizaciones de deseabilidad que se tienen.

2.0 Deseo y deseabilidad

Las opiniones de John Stuart Mill nos ofrecen un contraste instructivo. Aparentemente, Mill equiparaba desear un objeto con creerlo deseable, y sostenía en particular que «la única evidencia que es posible aducir de la deseabilidad de una cosa consiste en que la gente, realmente, la desea» (Mill 1910: 32). Esas afirmaciones son vagas por lo que toca a la clase de equivalencia en cuestión, por lo que toca a las conexiones, si las hay, entre la deseabilidad y el consenso de los deseos, y por lo que toca a las consecuencias para situaciones imaginables en las cuales los deseos de la «gente» difieren de lo que «realmente» son. Pero en este punto, lo que más interesa es, pese al desagrado de Mill ante las reflexiones sobre cualquier «tema de Ontología», el espíritu claramente reduccionista en que se apoyan sus opiniones: un espíritu que se dirige contra la predisposición de la humanidad «a creer que cualquier sentimiento subjetivo para el que no se tiene otra explicación, es una revelación de alguna realidad objetiva» (*ibid.*: 39). Siguiendo esa predisposición, pensaremos que todos nuestros deseos son revelaciones de deseabilidades objetivas; si nos oponemos a ella totalmente, pensaremos que los deseos «reales» de la «gente» son la fuente inmediata y última de todo discurso sobre la deseabilidad.

Que algo ha salido mal queda atestiguado por el hecho de que nadie se permitiría la supuesta predisposición con respecto a *todos* sus deseos. Y si se aceptara que la predisposición interviene más fácilmente en el caso en que formen parte de la subclase problemática de los deseos motivados, alguna explicación debe darse de este hecho. Sólo teniendo a mano esa explicación resultará posible determinar con exactitud hasta qué grado la teoría de Mill es revisionista. Sin embargo, liberada de todo espíritu reduccionista semejante (sin caer por ello en el «trascendentalismo»), puede enunciarse una afirmación de apariencia similar a la ofrecida por Mill que exprese la concepción presente del deseo y la deseabilidad en relación con la subclase problemática de los deseos motivados. Desear no es en este caso equivalente de «creer deseable» —de tener, por ejemplo, una caracterización de deseabilidad—, ni la noción de deseabilidad se entiende aquí en términos de lo que se desea «realmente». No obstante, la siguiente afirmación parece correcta para estos casos: la mejor evidencia que se puede aducir de la deseabilidad de una cosa consiste en una caracterización de deseabilidad de ella, una especificación de sus

características de deseabilidad externas; y la mejor evidencia de que se ha aducido la mejor evidencia se relaciona con la eficacia de esa caracterización de deseabilidad para provocar deseos.

2.10 Deseo y placer

Consideremos ahora la distinción que traza Stephen Schiffer entre lo que llama deseos «generadores de razones (g-r)» y deseos «derivados de razones (d-r)». Los segundos, nos dice Schiffer, son los mismos que los deseos «motivados» de Nagel, de modo que las observaciones de Schiffer sobre los deseos «derivados de razones» pueden contribuir de manera directa a elucidar la clase *general* de los deseos motivados. Por otra parte, como en breve veremos, sus observaciones sobre los deseos «generadores de razones» pueden ayudar a elucidar una *sub*clase de los deseos «no motivados» de Nagel. Schifer escribe:

> Si nuestro deseo de ϕ fuera un deseo d-r y si realmente hiciéramos ϕ, habría entonces una razón que sería tanto la razón del deseo de ϕ como la razón de que hiciéramos ϕ, y esta razón será enteramente independiente, lógicamente, del hecho de que deseamos ϕ. Pensamos que el acto ϕ es deseable de cierta manera, y es porque pensamos que el acto de ϕ es deseable de esa manera por lo que tanto deseamos ϕ como realizamos ϕ; no es que deseemos ϕ la razón por la que el acto ϕ nos parece deseable; cuando nuestro deseo es un deseo d-r, creemos que, incluso si no tuviéramos el deseo, tendríamos una razón para tenerlo [...] el punto final se encuentra, por así decirlo, no tanto en el deseo como en su objeto.
>
> Las cosas son exactamente al revés cuando pasamos a los deseos g-r. Cuando el deseo por el que actuamos es un deseo g-r de ϕ, la razón por la cual realizamos ϕ y, de modo típico, la única razón que tenemos para realizar ϕ, está dada enteramente por el propio deseo de ϕ [...]
>
> No es porque una cosa sea deseable de cierta manera por lo que tenemos un deseo g-r de ella; todo lo contrario: es deseable de esa manera precisamente porque tenemos el deseo [...]
>
> [C]uando actuamos con base en un deseo g-r, actuamos para la obtención de placer y el alivio de la incomodidad —habitualmente ambas cosas, pero siempre una o la otra—, que nuestra acción proporciona [...]
>
> De suerte que, por lo regular, con los deseos g-r, su condición *sine qua non* es que son deseos que, casi siempre, resulta placentero satisfacer e incómodo sobrellevar, y siempre lo uno o lo otro; de hecho, el placer y el alivio de la incomodidad anticipados están casi siempre inextricablemente relacionados, por cuanto lo que anticipamos es justamente el placentero alivio de la incomodidad [...] (Schiffer 1976: 197-198)

Si, digamos, el deseo del placer que proporcionará el acto de comer es un caso de deseo «derivado de una razón», entonces se pensará que el placer es independiente de que ese deseo tenga lugar. Ese razonamiento tiene que parecerle insensato a alguien que, como el obispo Butler, sostenga que todo placer depende de que el deseo alcance su objeto apropiado, y que todo placer emana del hecho de que el acto de ϕ satisface el deseo de ϕ. Pero esto es falso: imaginemos que, ocupada la mente en otro asunto, el olfato se ve repentinamente gratificado con el perfume de una rosa (véase Dent 1984: 39). El ejemplo pertenece a Burke, y es quizá mejor que el que comúnmente se prefiere del comer, ya que resulta difícil concebir un caso posible y convincente en el que uno *se sorprenda* comiendo; más aún, en el ejemplo de Burke es más fácil comprender la posibilidad de que, una vez que la experiencia ha ocurrido, se llegue a tener el deseo de que ese placer se repita, y en ello el placer anticipado es todavía claramente independiente de que se dé el deseo. Sin duda, la satisfacción de uno cualquiera de la mayoría de mis deseos me proporciona placer; pero ello puede ser, meramente, un agradable efecto secundario de la estricta satisfacción del deseo. Por lo tanto, en esta extensión del ejemplo de Burke, un encuentro casual con una rosa inadvertida podría proporcionar placer y satisfacer el deseo: pero la persona en cuestión debería admitir razonablemente la posibilidad de que tal encuentro le produjese tal placer aun cuando no hubiera tenido el deseo que en efecto se da.

El asunto es diferente si el deseo, por ejemplo, del placer que proporciona comer es un caso puro de un deseo «generador de razones». La persona quiere comer, digamos, debido tanto a una sensación incómoda de hambre como a una anticipación del placer que ello le producirá. En un caso semejante, el placer anticipado está condicionado por la satisfacción del deseo; al considerar una circunstancia posible en que careciera del deseo que realmente tiene, el agente estará considerando una posibilidad en la cual no existiría ese placer anticipado producido por el comer.

En nuestros términos, la diferencia más general entre estos dos tipos de casos se expresa con sencillez. Si un agente actúa conforme a un deseo «generador de razones» para obtener placer, una especificación completa de su caracterización de deseabilidad para la acción ejecutada tiene que hacer referencia al deseo que realmente tiene: la deseabilidad de la acción es consecuencia de que ese deseo se dé. Vale decir, una caracterización de deseabilidad semejante no será lógicamente independiente de una afirmación de la existencia del deseo. Pero no se da ninguna condición semejante respecto de las especificaciones completas de las caracteriza-

ciones de deseabilidad en los casos de deseos «derivados de razones»: en pocas palabras, en tales casos la deseabilidad de la acción ejecutada es independiente de la existencia real del deseo (véase Platts 1980b: 77).

Esa explicación de la diferencia tiene que entenderse a la luz de otra. En el caso de que el deseo fuera todavía un deseo «derivado de razones» (es decir, motivado), al contemplar una posibilidad en la que careciese del deseo que de hecho tiene realmente, el agente considerará, aquí y ahora, que la deseabilidad de que ejecute una acción del tipo que de hecho ejecuta no se ve afectada por la supuesta falta de deseo por su parte en el mundo que imaginó; mientras que en el caso de que su deseo real fuera un puro deseo «generador de razones», pensará, aquí y ahora, que esa deseabilidad ha quedado eliminada en las circunstancias imaginadas. Por último, adviértase que, en los casos de deseos «generadores de razones», resultaría un error obvio tratar de aplicar directamente el hablar de que un deseo es una respuesta razonable a las características de deseabilidad «externas» de la acción en cuestión, pues esas características no existen independientemente de que el agente tenga realmente tal deseo.

2.11 El simple querer

Hasta el momento hemos distinguido tres categorías de deseos. En la terminología de Stephen Schiffer existe la categoría de los deseos «generadores de razones». Éstos tienen un carácter esencialmente fenomenológico, pues suscitan una sensación desagradable cuando no se los satisface, agradable o placentera al ser satisfechos, o ambas cosas. Desde un punto de vista más teórico, un deseo de esta índole es lógicamente autorreferencial en el sentido de que toda especificación completa del objeto del deseo —toda caracterización de deseabilidad completa de su objeto— hace referencia lógica a la existencia del deseo mismo. Más aún, el modo como se efectúa la referencia sirve para explicar por qué la «deseabilidad» del objeto del deseo no «trasciende» la existencia de dicho deseo: al contemplar un mundo posible en el cual careciese del deseo en cuestión, el agente no vería nada deseable en la realización ahí de lo que es su objeto de deseo real según se especifica en alguna caracterización de deseabilidad *incompleta* del mismo.

En segundo lugar, dentro de la categoría más amplia de los deseos «motivados», de los deseos «derivados de razones», hay una subcategoría caracterizada en los términos siguientes: para cualquier deseo de este tipo, una caracterización de deseabilidad completa de su objeto hará referencia, lógicamente, no a la existencia de ese mismo deseo, sino más

bien a la existencia de algún otro deseo que el agente tenga, sea cual sea el tipo de este otro deseo. Resulta obvio, por tanto, que, al menos en términos generales, mientras que la presunta deseabilidad del objeto de un deseo semejante puede «trascender» la existencia de ese mismo deseo, no «trascenderá» la existencia del otro deseo que tiene el agente. Más aún, en cualquier caso particular de este tipo, el agente puede fundamentar la afirmación de que la caracterización de deseabilidad en cuestión sirve para captar la deseabilidad del objeto, haciendo referencia tanto a la existencia de su otro deseo como a los hechos pertinentes sobre el papel que desempeña el objeto así caracterizado en el proyecto de satisfacer ese otro deseo.

Finalmente, de las categorías de deseos que hemos distinguido hasta ahora, nos queda una por analizar: la problemática subclase de los deseos «motivados», en la que se centran las discusiones de Nagel y McDowell. Un deseo de este tipo no posee un carácter fenomenológico esencial y la caracterización de deseabilidad completa de su objeto no necesita hacer referencia lógicamente a la existencia del deseo mismo. Pero tampoco es necesario que una caracterización de deseabilidad completa del objeto de un deseo de este tipo haga referencia lógicamente a la existencia de algún otro deseo que el agente tenga. Sin embargo, en términos más positivos, para cualquier deseo de este tipo, se cuenta con la posibilidad de que, dentro de un contexto específico, el agente aduzca alguna razón que apoye la tesis de que la caracterización de deseabilidad completa de su objeto es verdaderamente de tal índole que sirve para captar la deseabilidad de ese objeto; el agente no está destinado a conformarse con la idea de que simplemente lo quiere.

Pero queda una última cuestión que viene a complicar el cuadro. Schiffer nos dice que los deseos «motivados» de Nagel son los mismos que sus deseos «derivados de razones», mientras que sus deseos «generadores de razones» sólo constituyen un subconjunto propio de los deseos «no motivados» de Nagel (Schiffer 1976: 198–199, n. 4). ¿Cuáles son, entonces, los deseos «no motivados» que, sin embargo, no son, en términos de Schiffer, deseos «generadores de razones»? La caracterización general de los deseos de esta clase restante es esencialmente negativa. Los deseos en cuestión carecen del carácter fenomenológico esencial de los que son «generadores de razones», y una presunta caracterización de deseabilidad del objeto de un deseo de esa índole no necesita hacer referencia lógica a la existencia de ese deseo (de este modo es posible que la presunta deseabilidad de ese objeto «trascienda» la existencia del deseo

mismo). Pero tampoco sucede que, para cada deseo de este tipo, una presunta caracterización de deseabilidad tenga que hacer referencia lógica a la existencia de algún otro deseo que el agente tenga. Y, por último, en el caso de un deseo de este tipo, no hay tampoco ninguna posibilidad de que el agente aduzca alguna razón para fundamentar la afirmación de que la presunta caracterización de deseabilidad es verdaderamente tal que sirve para captar la deseabilidad del objeto. Estos deseos son los que más invitan a pensar: *simplemente lo quiere.*

Es en estos casos donde podríamos seguir con más naturalidad a Epicteto cuando dice que lo que se desea, se desea por azar. Nagel mencionaba algunos casos que podrían parecer buenos candidatos en este aspecto: que hoy elija comer caquis, el anhelo de que mis cenizas sean esparcidas sobre el Potomac, mi deseo de que haya perejil en la Luna (Nagel 1970: 42–45). Pero no tenemos razón alguna para restringir los posibles candidatos a elecciones, caprichos y antojos de ese tipo: ¿por qué no incluir entre ellos el deseo de bienestar para otra persona, o incluso el deseo de acabar con el sufrimiento y proteger a los débiles? Esta subcategoría distintiva de deseos no motivados no tiene que limitarse a lo caprichoso, lo gracioso y lo frívolo, ni nadie que dedique su vida a actuar en conformidad con modelos pertenecientes a esta subcategoría tiene que convertirse por ello en un personaje salido de La Novela Moderna (no californiana, desde luego). Pero debe subrayarse que los candidatos mencionados no son más que eso: muy pocas especificaciones cotidianas, prefilosóficas, del contenido de un deseo sirven para determinar el sitio que ese deseo ocupa dentro de la taxonomía que aquí se presenta. Los ejemplos que ofrece Nagel parecen buenos candidatos a ser miembros de este último tipo de deseo, especialmente el de los caquis, que evoca la experiencia común de escoger de un menú al azar; pero la cuestión de si ha de elegirse una ejemplificación específica de alguno de esos candidatos revertirá en la cuestión de si tal ejemplificación satisface la caracterización dada de este último tipo de deseos.

En cierta ocasión, Bernard Williams pensó en un caso que tenía la siguiente estructura: el estado de cosas deseado *S* no contiene al deseador, ni una especificación de ese estado que haga referencia esencial a él, o que implique su existencia o alguna relación entre él y algo o alguien que se halle en *S*; mucho menos se requiere una especificación de *S*, el estado que satisface el deseo, que mencione el deseo mismo. Entre los ejemplos que da Nagel, sólo el del perejil en la Luna tiene probabilidades de pertenecer a este tipo; pero los deseos de que María se encuentre bien o de

que se acabe con el sufrimiento y se proteja a los débiles también podrían ser de este tipo. Acerca de su ejemplo, Williams señaló lo siguiente:

> Siendo esto así, *S* podría darse sin que nuestro hombre existiera en absoluto, y éste debería ser capaz de reflexionar al respecto. Lo podemos invitar a considerar dos posibilidades para un mundo en que él no exista, una en la cual *S* se produce y otra en la que no se produce [...]
>
> De lo que ya hemos expuesto, debe seguirse que nuestro hombre prefiere una de estas posibilidades sobre la otra: prefiere, desde luego, la que contiene a *S*. (Williams 1973: 264)

Supongamos ahora que, dentro de los términos de la presente taxonomía, el caso descrito es un ejemplo del último tipo de deseos que acabamos de distinguir. Este caso sirve entonces para mostrar que la «trascendencia» de la presunta deseabilidad del objeto del deseo frente a la existencia de ese mismo deseo, no sirve para marcar la distinción entre los deseos motivados y los no motivados. Esa «trascendencia» sirve, en general, sólo para diferenciar a otros deseos de los deseos «generadores de razones».

Supongamos que tenemos a un agente en el mundo real *m*, con cierto conjunto de deseos, e imaginemos que el agente contempla un mundo imaginario *m'* en el cual posee un conjunto distinto de deseos. Se plantean entonces las siguientes preguntas:

(1) ¿Qué quiere el agente, aquí y ahora, en el mundo real *m*, que haya sucedido en el mundo *m'*?
(2) En el mundo imaginario *m'*, ¿qué querría, allí y entonces, que hubiera sucedido en ese mundo?
(3) Dado su conjunto real de deseos en *m*, ¿qué dice él, aquí y ahora, de su conjunto de deseos en el mundo imaginario *m'*, y por ende del contenido de su respuesta a (2) en este mundo?

Al restringirse a casos de mundos sin un agente, Williams se limita, desde luego, a la primera de estas preguntas: su respuesta a su pregunta es evidentemente correcta. Pero la pregunta que nos interesa para nuestros propósitos es la tercera.

Supongamos que el conjunto pertinente de deseos poseído por el agente en el mundo real pertenece al tipo recién reconocido: ¿qué puede decir él aquí de un mundo imaginario en el cual sus deseos difieren en cuanto al asunto pertinente? Allá se ve atraído por el huachinango, no por los caquis; por el ajo, no por el perejil; por el Támesis, no por el Potomac; por Juan, no por María; por el desarrollo de algún ideal individual, no por la

supresión del sufrimiento y la protección de los débiles. Puede decir que sus deseos en ese mundo imaginario difieren de como son en el mundo real y, por lo tanto, puede decir (por lo menos en algunas circunstancias concebibles) que la probabilidad de que en ese mundo imaginario suceda lo que desea en realidad aquí y ahora se ha reducido. Pero eso no necesita ser una gran crítica del mundo imaginario, así como tampoco es una gran crítica del mundo real el hecho de que no se cumplan en él muchos de mis deseos. Tampoco cambiaría sustancialmente el caso si al conjunto de deseos reales del agente se agregara algún deseo «de orden superior» *del mismo tipo*, conforme al cual poseería sus (otros) deseos en todos los mundos imaginarios. Esto multiplica las diferencias entre los mundos y garantiza que algunos de sus deseos reales no se cumplan, *ex hypothesi*, en el mundo imaginario; pero eso no pondría de manifiesto que este último, en algún otro sentido, estuviera peor o fuera menos deseable.

Obsérvese cuán distinto sería el asunto si el conjunto de deseos pertinente en el mundo real estuviera compuesto por deseos motivados. Tales deseos exigen que el agente piense que el objeto del deseo es deseable independientemente de que *él* tenga *ahora* el deseo; lo cual requiere a su vez el pensamiento de que, *ceteris paribus*, es deseable que tenga ese deseo y que lo siga teniendo en los mundos imaginarios. Pero *esa* deseabilidad no necesariamente es mera consecuencia de que el agente tenga otros deseos de la última clase reconocida. Más aún, con los deseos motivados existe la posibilidad de que el agente aduzca razones en apoyo de la caracterización de deseabilidad pertinente, y ello trae consigo en este caso la posibilidad de una crítica sustantiva, en la respuesta del agente a la pregunta (3), de la parte que desempeña él en el mundo imaginario.

El objetivo no es negar la existencia de los deseos de la última clase reconocida ni condenarlos. Más bien se trata de enfatizar el carácter distintivo de los deseos motivados, no en lo que se refiere a la «trascendencia», sino en cuanto a la posibilidad de una discusión y evaluación de esos deseos que invoque razones.

La concepción de Schiffer de que los deseos «generadores de razones» son autorreferenciales está fundada en una propuesta de forma lógica para los enunciados que los atribuyen. Sería un error propugnar el uso de esa propuesta para todos los miembros de la última clase de deseos reconocida; pero en un nivel menos formal, hay un argumento a favor de una concepción similar de tales deseos. Ante cualquier caso aparente de deseo, podemos intentar determinar alguna caracterización de deseabilidad de su objeto que el deseante posea o de la cual pueda disponer.

Pero, ¿qué puede ofrecerse en el caso de un deseo de esta última clase reconocida? El deseante podría decir: «Eso es exactamente lo que quiero»; o podría agregar: «Y ustedes también lo quieren»; o podría agregar además: «Y ellos también lo quieren». No podría decirse nada más que no equivaliera a «convertir» el deseo en cuestión en uno de la clase de deseos «generadores de razones» o en uno de la de los deseos motivados. Así pues, para el agente la pregunta de *por qué* quiere el objeto de que se trata tendrá que ser contestada como: *es solamente lo que quiere*; o como: *es solamente lo que nosotros (él y su público) queremos*; o: *es solamente lo que ellos (él y otros) quieren.* El agente no tiene a su disposición nada más que una «pura referencia» a algunos ejemplos del tipo de deseo —incluyendo, por así decirlo, su propio ejemplo—. Y por ello hay tal escasez de discusiones razonables cuando se contemplan las variaciones posibles de tales deseos.

Quizá mi vida filosófica se facilitaría si se pudiera afirmar de un modo razonable que los deseos de esta clase recién reconocida son, a lo sumo, casos limítrofes de deseos; es decir, si pudiera afirmarse razonablemente que el *solo* querer es *apenas querer.* Sin embargo, esto no es posible, y el hecho de que no lo sea tiene además cierta importancia. Páginas antes (pp. 41–42), justa o injustamente, le adjudiqué a Anscombe la condición siguiente sobre las atribuciones de deseos: una condición necesaria de la aceptabilidad de la atribución de un deseo específico a un agente es que quien lo atribuye haya «alcanzado y vuelto inteligible» alguna «caracterización de deseabilidad» del estado de cosas (potencial) especificado por el contenido del deseo así atribuido. Y la noción de caracterización de deseabilidad que allí se invocaba se explicaba de la siguiente manera: tal caracterización del objeto del deseo es la adecuada para comunicar la deseabilidad de dicho objeto, la que sirve para llevar a buen término el cuestionamiento de por qué el deseante desea ese objeto poniendo de manifiesto, a los ojos de los demás, el bien imaginado que encierra. Pero resulta claro que en los casos de deseos pertenecientes a la última clase reconocida no se dispone de tal caracterización de deseabilidad. Ni la noción requerida de *comunicar la deseabilidad* (en tanto opuesta a la mera «infección» por un deseo) ni la noción requerida de *acuerdo* (en tanto opuesta a la mera coincidencia) tienen aplicación alguna en estos casos. Sin embargo, también está claro que podemos atribuir y atribuimos deseos de este tipo; que en tales casos podemos reconocer y reconocemos *el objeto* del deseo sin pensar en *su deseabilidad.* Admitimos que otras personas pueden desear toda clase de cosas hacia las cuales nosotros no sentimos la menor inclinación de desear, aunque también nos conside-

ramos capaces de indicar cuáles son esas cosas. No conociendo ninguna razón para criticar esa idea o esas atribuciones, concluyo que los deseos pertenecientes a esta clase recién reconocida nos exigen *o bien* rechazar la condición universal atribuida a Anscombe sobre las atribuciones de deseos, *o bien* reinterpretar dicha condición para que invoque una noción más débil de «caracterización de deseabilidad». Puesto que el segundo camino amenaza con volver casi trivial la condición respectiva —si la noción de «caracterización de deseabilidad» se interpreta siguiendo los lineamientos, digamos, de una caracterización adecuada para comunicar *el objeto* del deseo—, prefiero el primero. Pero esto no exige el rechazo del ideal regulativo según el cual debemos buscar, por lo menos en general, caracterizaciones de deseabilidad de los objetos de los deseos, ni exige la negación de la importancia que para la comprensión general del concepto de deseo tiene la noción de una caracterización de deseabilidad.

Los deseos pertenecientes a la última clase reconocida son, pues, casos de deseos que carecen de (la posibilidad de) caracterizaciones de deseabilidad. Esto obliga a una serie de breves comentarios finales. (a) Esta idea se insinuó en la discusión inicial sobre estos deseos (pp. 77–78) mediante el uso repetido de la palabra «presunta»; de aquí en adelante se usará la expresión «caracterizaciones de deseabilidad*» cuando se esté hablando de tales casos, lo que habrá que entender en términos de una caracterización adecuada para comunicar sólo *el objeto* del deseo, y lo mismo vale para las expresiones emparentadas. (b) Dada la estipulación hecha en las páginas 61–62 de que la noción de «percepción de deseabilidad» debe ser comprendida en vinculación con la de caracterización de deseabilidad disponible, se sigue que los deseos pertenecientes a esta última clase reconocida son casos de deseos sin percepciones de deseabilidad. Creo que esta consecuencia será bienvenida si mantenemos la seriedad del uso del modelo *perceptual* para los deseos motivados. (c) Por lo menos aquellos casos de actividades de sustitución que realmente contradicen la identificación contemplada de desear con creer deseable (pp. 61–62) serán casos en los cuales los deseos operativos sean deseos de esta última clase reconocida. (d) Resulta ahora claro que cualquier intento de identificar el desear con el «creer deseable» afronta una dificultad con respecto a la manera de elucidar esta segunda noción de tal suerte que la pretendida identificación no sea ni obviamente falsa ni absolutamente trivial (véanse las pp. 54–58, 61–62). (e) Es en los casos de la última clase de deseos reconocida, en la que se incluyen los deseos puestos de manifiesto en las actividades de sustitución pertinentes, en los que, conceptualmente, las

conexiones entre desear y tratar de conseguir (o tener como meta) son más estrechas; en relación con esto, es en los casos de deseos pertenecientes a la última clase reconocida, y *no* en los casos de los deseos motivados, en los que en cierto sentido sólo encontramos una dirección motivacional y *nada más* (véanse las pp. 62–63, 65, 69). (f) Recurrir a ejemplares de esta última clase de deseos reconocida podría servir para construir una nueva variación del caso de Sísifo (pp. 42–46). Pero la no aplicabilidad, a tales ejemplares, de las nociones de caracterización de deseabilidad y percepción de deseabilidad, junto con otras consideraciones que se aducirán en el próximo capítulo, dan suficientes razones para dudar del interés que tendría ese caso con vistas a iluminar problemas de significado y valor.

Habiendo rechazado en el primer capítulo la falsa concepción clásica del deseo, éste comenzó con el planteamiento de una alternativa: dicho brevemente, la que sostiene que desear debe equipararse con creer deseable. La alternativa planteada afronta tres dificultades: la existencia de lo que se denominó «actividades de sustitución» amenaza con volverla falsa o trivial (pp. 61–62); esta alternativa no toma en cuenta la distintiva dirección de adecuación de los deseos frente al mundo (p. 69); y los casos de «simple querer» extienden a la sustancia de esa concepción la amenaza que de todas maneras plantea la clase más restringida de las actividades de sustitución (pp. 82–83). No obstante, el planteamiento de esta falsa concepción alternativa sirvió para introducir varias ideas que habrían de ser útiles más adelante; en este sentido, la falsa concepción alternativa resulta más instructiva que la clásica.

Puesto que una concepción abstracta del deseo en cuanto a la dirección de adecuación distintiva es *en el mejor de los casos* un tanto endeble, me vi obligado a desarrollar una descripción detallada de una posible taxonomía de los deseos. (Según esa taxonomía, es plausible la conjetura de que los partidarios de la falsa concepción clásica se han concentrado exclusivamente en los deseos «generadores de razones», mientras que los partidarios de la falsa concepción alternativa que se examinó se han concentrado primariamente en los deseos «derivados de razones» o «motivados».) El resultado no es, en realidad, más que una posible taxonomía de los deseos, y la descripción exacta que de ella se dio no es más que una de sus posibles descripciones. El mérito de esa descripción de la taxonomía en cuestión tiene que evaluarse a la luz de los resultados de su empleo subsecuente.

Spinoza sostenía, en efecto, que el subjetivismo sobre los valores es consecuencia de una concepción del deseo como poder activo; y de hecho

existe una coincidencia histórica sorprendente entre ese subjetivismo y esa concepción. He rechazado esa concepción, pero también he señalado que es improbable que un tratamiento abstracto de un concepto tan general como el de deseo nos permita avanzar mucho en el camino hacia una concepción sustancial. Lo sustancial comienza con el reconocimiento de la diversidad de los tipos de deseo. El capítulo siguiente se ocupa de las consecuencias que tiene para la teoría del valor el reconocimiento de esa diversidad.

Capítulo 3

VALORES

> Troilo: ¿Qué otro valor tiene un objeto además del que se le da?
> Héctor: Pero la valía de un objeto no depende de una apreciación individual; su mérito y su importancia provienen tanto de su precio intrínseco como de la estimación del tasador; hacer el culto más grande que el dios es loca idolatría, y la pasión delira cuando atribuye cualidades de que es fanática a un objeto que no tiene ni sombra de ese mérito apreciado.
>
> William Shakespeare

> Como profesor universitario, busco un ambiente de estudio neutral, libre de valores, con el fin de alcanzar el análisis y la objetividad académicos. Está dentro de la autoridad, la discreción y la libertad académica de un profesor imponer condiciones apropiadas y razonables. Algún grado de autoridad es inherente a la relación entre maestro y estudiante. Yo impongo ciertas condiciones: decencia, no fumar, no interrumpir, no cometer abusos personales, no proferir estribillos racistas o antisemitas, no exhibir fotos de políticos, no portar escarapelas, no mostrar insignias políticas de ninguna especie, sean de izquierda o de derecha.
>
> Alec Samuels en una carta a *The Guardian*, 1983

3.1 Un escepticismo cauteloso

Philippa Foot manifestó una vez un escepticismo cauteloso en los siguientes términos:

> Supondría, por ejemplo, que en ciertas investigaciones morales fundamentales, podríamos encontrarnos apelando al hecho de que la vida humana es un valor. Pero, ¿entendemos realmente esta idea? ¿Sabemos lo que queremos decir cuando hablamos de que *alguna cosa* tiene valor, o incluso de que la valoramos, en tanto opuesto a quererla o a estar dispuesto a tomarse la molestia de conseguirla? No sé de ningún filósofo, vivo o muerto, que haya podido explicar esta idea. (Foot 1978a: 165)

Tal escepticismo es ciertamente cauteloso, pues está claro que Foot cree que aquí hace falta mucho trabajo si queremos tener alguna esperanza de comprender en detalle lo que sucede en cualquier discusión auténtica sobre moralidades divergentes:

> [P]odemos, por ejemplo, entender a un hombre que, al término de su vida, dice que ha desperdiciado el tiempo en «cosas sin importancia». Pero, ¿cuáles son las cosas «importantes» si no las cosas triviales a las que dedicamos tanto tiempo? Obviamente, tales preguntas son pertinentes en discusiones fundamentales sobre la moralidad de otras sociedades y de la nuestra. Es imposible juzgar la moralidad de una sociedad si no podemos hablar de sus valores, y tenemos que poder manejar la idea de los falsos valores si pretendemos establecer qué está mal en una sociedad materialista como la nuestra. Pero, ¿en qué consiste tener falsos valores si no en estimar demasiado cosas que no importan mucho? (*Ibid.*: 166)

Así pues, el escepticismo no parece menos importante que cauteloso.

Un examen paciente de las diversas formas en que se expresan nuestros pensamientos sobre el valor —de las variedades de construcciones valorativas que empleamos y de los diferentes contextos en que naturalmente las empleamos— no carecería de interés ni resultaría, según sospecho, infructuoso. Pero para nuestros propósitos presentes podría bastar con distinguir de un modo muy tosco y apresurado algunas de esas construcciones y algunos de sus contextos naturales de uso.

El contexto de uso más general en que me concentraré inicialmente es aquel en el que se despliegan construcciones valorativas como parte del intento de explicar acciones intencionales. Y aquí es preciso introducir una distinción tosca. Por un lado, podríamos decir, como parte de una presunta explicación de una acción dada, que el agente hizo lo que hizo porque valora, digamos, la honestidad, o porque concede valor a los tratos honestos, o porque aprecia la honestidad, o porque cree que la honestidad es importante, o porque cree que la honestidad es valiosa o tiene valor. Por otro lado, dentro del mismo contexto, nosotros podemos decir que el agente hizo lo que hizo porque reconoce el valor de la honestidad, o porque reconoce que la honestidad es importante, o porque reconoce que la honestidad es valiosa o tiene valor, o incluso debido al valor de la honestidad, o debido a que la honestidad es importante, o debido a que la honestidad es valiosa o tiene valor. El fundamento de esta tosca distinción resulta obvio: mientras que el primer grupo de construcciones valorativas son todas estricta y literalmente neutrales con respecto a

la cuestión de la actitud asumida hacia la honestidad por la persona que expresa la presunta explicación, el uso de cualquier construcción valorativa del segundo grupo, dentro de este contexto, requiere por lo menos que la persona que enuncia la explicación valore la honestidad. Puede detectarse otra distinción tosca similar en las presuntas explicaciones de acciones intencionales en las que se hace referencia a lo que podría denominarse *evaluaciones comparativas resultantes de valoraciones*; como, por ejemplo, cuando decimos, con respecto a un examinador de filosofía, que otorga una calificación más alta a un estudiante que a otro porque él, el examinador, valora la claridad y estima la precisión, o porque reconoce que la claridad es importante y la precisión es valiosa. El «compromiso valorativo» de la persona que expresa la explicación es a todas luces más débil cuando la explicación enunciada es del primer tipo. Para nuestros fines actuales, no hay necesidad de negar la aparente posibilidad de que, digamos, un examinador de filosofía que desde hace tiempo ha dejado de preocuparse por la claridad y la precisión pudiera continuar siendo capaz de calificar o clasificar adecuadamente los trabajos de los estudiantes en cuanto a su claridad y precisión. No obstante, como mera estipulación terminológica diré que, en esta posibilidad, ni la calificación ni la clasificación expresan una *evaluación*. Según el uso que aquí le doy, la evaluación es siempre resultado de una valoración.

3.2 Del querer al valorar

El escepticismo cauteloso de Foot era acerca de lo que se podría querer decir «cuando se habla de que *alguna cosa* tiene valor, o incluso de que la valoramos, en tanto opuesto a quererla o a estar dispuesto a tomarse la molestia de conseguirla». Pero ese escepticismo acerca de que exista alguna teoría del valor y de la valoración fuera de la que se encuentra fundada en una *identificación* del valorar con el querer o con el estar dispuesto a tomarse la molestia de conseguir, no es todavía suficientemente cautelosa. Acabamos de indicar algunas distinciones intuitivas, preteóricas, en los modos en que hablamos naturalmente acerca del valor y de la valoración. Si ahora colocamos esas distinciones preteóricas junto a las distinciones establecidas en el capítulo anterior entre las cuatro clases de deseos, observaremos ciertas dificultades en el camino de la simple identificación que contempla el escepticismo de Foot:

(1) Dentro de la clase de los «deseos no motivados» de Thomas Nagel distinguimos una subclase que, siguiendo a Stephen Schiffer, llamamos

«deseos generadores de razones». Tales deseos tienen un carácter esencialmente fenomenológico por cuanto producen una sensación desagradable cuando no se los satisface, o una sensación agradable o placentera cuando se los satisface, o ambas. Sin embargo, en un nivel más teórico, cualquiera de esos deseos es lógicamente autorreferencial, en el sentido de que toda especificación completa del objeto del deseo —toda caracterización de deseabilidad completa de su objeto— hace referencia lógica a la existencia del deseo mismo. Más aún, la forma en que dicha referencia se lleva a cabo sirve para explicar por qué esta deseabilidad del objeto del deseo no «trasciende» la existencia de ese deseo: al contemplar un mundo posible en el que carezca del deseo en cuestión, el agente no habría de encontrar ninguna deseabilidad de este tipo en la realización, en ese mundo, de lo que constituye su verdadero objeto de deseo, según lo especifica cierta caracterización de deseabilidad incompleta de él. Los casos más conocidos de deseos pertenecientes a esta clase son los apetitivos, pero éstos no agotan en modo alguno los ejemplares de este tipo.

En los casos de deseos puramente de este tipo, sin duda es importante observar que hablar de valor o de valoración no cabe de manera directa y natural dentro de nuestra consideración de ellos —ni siquiera cuando el agente en cuestión está dispuesto a tomarse algunas molestias considerables a fin de conseguir el objeto de su deseo—. Supongamos que con frecuencia me invade una sed intensa y distintiva por cierto fluido específico, e incluso que de hecho paso gran parte de mi vida persiguiendo y buscando ese fluido. Aun así, resultaría extraño decir que yo valoro el fluido en cuestión, que le confiero valor o que creo que es valioso para mí; y sería aun más extraño recurrir de un modo directo, en cualquier caso que sea puramente de este tipo, a un discurso más impersonal acerca del valor. Toda identificación simple del valorar con el querer requerirá correcciones para tomar esto en cuenta.

Deben mencionarse dos complicaciones a fin de aislar la dificultad presente. (a) El discurso valorativo podría intervenir indirectamente en cualquier caso de este tipo, en el cual la búsqueda constante del objeto del deseo lleve al agente en cuestión a descuidar otras valoraciones, sean éstas suyas o de la persona que enuncia la explicación del caso del agente. (b) El discurso valorativo de un tipo u otro puede intervenir en casos de aspecto similar, en los cuales se recurre legítimamente, para dar razón de esos casos, a la idea de que, digamos, el agente en cuestión cree o sabe que el fluido en cuestión es necesario para su supervivencia. Pero si se recurre legítimamente a esa idea y ésta sirve para darle un lugar directo

al discurso valorativo de que se trata, esto se limita a mostrar que el caso en cuestión no es el de un deseo puro del tipo de los «generadores de razones», sino que más bien posee un carácter «híbrido».

La otra subclase de «deseos no motivados», el tipo (2), se caracterizó en términos negativos. Los deseos en cuestión carecen del carácter fenomenológico esencial de los que son «generadores de razones»; además, una caracterización de deseabilidad* completa del objeto de cualquiera de estos deseos no necesita hacer referencia lógica a la existencia de ese deseo, lo cual hace posible que la presunta deseabilidad* de ese objeto «trascienda» la existencia del deseo mismo. (Para la explicación de los asteriscos, véase la p. 82.) Pero tampoco sucede que, para un deseo de esta clase, una caracterización de deseabilidad* completa de su objeto tenga que hacer referencia lógica a la existencia de algún otro deseo del agente. Por último, no hay tampoco ninguna posibilidad, en el caso de un deseo de este tipo, de que el agente ofrezca una razón en apoyo de la afirmación de que la caracterización de deseabilidad* completa es tal que verdaderamente sirve para captar la deseabilidad del objeto. Tales deseos son los que más propician el pensamiento de que *simplemente lo quiere.*

Se ha reconocido aquí que hay deseos de esta clase, y se ha aceptado también que sus objetos no necesitan ser triviales ni caprichosos; debe admitirse ahora, además, que un agente puede estructurar gran parte de su vida alrededor de la consecución del objeto de tal deseo, y tomarse molestias considerables en su esfuerzo por satisfacer tal deseo. Sin embargo, aun cuando el objeto de algún deseo de este tipo no sea trivial y las molestias que el agente en cuestión se tome sean grandes, no por ello caben de manera natural, para dar razón de tales casos, discursos valorativos que no sean estrictamente «personales»: decir que el agente *valora* o *aprecia* el objeto de su deseo, hablar tal vez de que *confiere valor* a ese objeto, o quizás incluso de que el objeto *es importante para él.* Ningún discurso valorativo de carácter «impersonal» encuentra un lugar directo y natural dentro de la explicación de esos casos. Y la razón de ello no es difícil de descubrir: la razón que se refiere al rasgo muy distintivo de los deseos de este tipo, el rasgo de que no puede decirse nada en apoyo de la afirmación de que la caracterización de deseabilidad* presunta del objeto del deseo sirve para captar la deseabilidad de ese objeto. El asunto comienza y termina con (i) el hecho de que el agente quiera de esta manera y esté dispuesto a tomarse molestias considerables para conseguir el objeto no trivial de que se trate, y (ii) su valoración de ese objeto. O, tal vez mejor, el asunto comienza y termina con el hecho de que el agente quiera

de esta manera ese objeto no trivial, y tal manera de querer tal objeto *es* una forma de valorar tal objeto. La conclusión tentativa consiste, pues, en que una vez que se introducen las restricciones acerca del carácter no trivial del objeto de deseo y acerca de la disposición del agente a dedicar esfuerzos considerables en su consecución, la identificación contemplada en el escepticismo cauteloso de Foot es, con respecto a los casos de deseos de este segundo tipo, sustancialmente correcta.

Dentro de la clase de los «deseos motivados» de Thomas Nagel, distinguimos una subclase (3) que se caracterizó en los siguientes términos: para cualquier deseo de esta subclase, una caracterización de deseabilidad completa de su objeto hará referencia lógicamente no a la existencia del deseo mismo, sino más bien a la existencia de algún otro deseo que tenga el agente, cualquiera que sea la clase de este otro deseo. Así, mientras que la presunta deseabilidad del objeto de tal deseo puede «trascender» la existencia de ese mismo deseo, no «trascenderá» la existencia del otro deseo del agente, al cual se hace referencia lógica dentro de la caracterización de deseabilidad completa en cuestión. (Esta última afirmación tiene que restringirse para tomar en cuenta ciertos casos en los cuales el otro deseo del agente cae dentro de la subclase restante de deseos (4).) Además, en cualquier caso particular de este tipo, el agente puede apoyar la afirmación de que la caracterización de deseabilidad en cuestión sirve para captar la deseabilidad del objeto mediante la referencia tanto a la existencia de su otro deseo como a los hechos pertinentes acerca del papel que el objeto así caracterizado desempeña en el proyecto de satisfacer ese otro deseo. Dos ejemplos aparentemente obvios de deseos de esta tercera clase son, primero, aquellos en los que el objeto de deseo es un medio para alcanzar el objeto que por otro lado se desea, y, segundo, aquellos en los que el objeto de deseo forma parte constitutiva del objeto que por otro lado se desea.

En un caso cualquiera de deseo de esta tercera clase, resulta perfectamente natural decir que el agente valora o confiere valor al objeto de su deseo, y en ciertos casos sería más o menos natural afirmar que el objeto del deseo es de, o posee, un valor instrumental o constitutivo para el agente en cuestión. Pienso que incluso sería más o menos natural hablar, en tales casos, de una manera todavía más «impersonal»: o sea, decir que el objeto del deseo posee un valor instrumental o constitutivo adherido a él —y adherido a él no estrictamente por el agente, sino más bien por el otro deseo que tiene el agente—. Este paso a lo «impersonal» tiene cierta significación: por así decirlo, dada la existencia del otro deseo del agente

junto con las consideraciones fácticas pertinentes, instrumentales o constitutivas, el objeto instrumental o constitutivo posee valor instrumental o constitutivo para el agente, tenga éste o no el deseo correspondiente de esta tercera clase. En tal caso, y todavía por así decirlo, el objeto instrumental o constitutivo es por lo menos digno de ser deseado por el agente, sería por lo menos un objeto razonable de deseo para el agente, tenga éste de hecho o no el deseo correspondiente. Más aún, si en ese caso la reflexión por parte del agente acerca de las circunstancias da origen al deseo del que antes carecía, entonces ese deseo puede entenderse por lo menos como una respuesta razonable al hecho de que su objeto posee valor instrumental o constitutivo para el agente. El agente llega a valorar el objeto en cuestión porque reconoce que, dadas las circunstancias, ese objeto tiene valor para él. Y, finalmente, obsérvese que la descripción de ese tipo de casos en que interviene la reflexión puede aplicarse también a otros casos similares salvo por el hecho de que no hay razonamiento por parte del agente: la posibilidad de que el agente invoque, razonablemente, razones para desear un objeto no implica que la formación de ese deseo haya sido resultado de la reflexión.

Si lo que se acaba de decir es correcto, resulta obvio que cualquier propuesta que intente identificar la idea de que algo posee valor para un agente con la idea de que el agente lo quiere o está dispuesto a tomarse la molestia de conseguirlo, requerirá adaptaciones considerables.

(4) El último tipo de deseo distinguido fue el ejemplificado por los miembros de la subclase restante de los «deseos motivados» de Nagel, subclase en la que se centraban tanto la discusión de Nagel como la de John McDowell. Un deseo de esta especie no posee un carácter fenomenológico esencial, y una caracterización de deseabilidad completa de su objeto no necesita hacer referencia lógica a la existencia del deseo mismo: de ahí el contraste con los «deseos generadores de razones» del tipo (1). Pero una caracterización de deseabilidad completa del objeto de un deseo de este último tipo tampoco necesita hacer referencia lógica a la existencia de otro deseo del agente: de ahí su contraste con los deseos del tipo (3). Sin embargo, en términos más positivos, para cualquier deseo de este último tipo existe la posibilidad de que el agente aduzca, dentro de un contexto específico, alguna razón en apoyo de la afirmación de que la caracterización de deseabilidad completa de su objeto es verdaderamente de tal índole que capta la deseabilidad de ese objeto: el agente no está destinado a quedarse con la idea de que simplemente lo quiere. De ahí el contraste con los deseos del tipo (2).

Es en los casos de acciones fundadas en deseos de este tipo donde resulta natural que los discursos valorativos más completamente «impersonales» vengan a ocupar un sitio dentro de nuestras explicaciones de los mismos. El agente hace lo que hace porque cree que algo —la honestidad o la verdad científica, por ejemplo— es un valor, o tiene valor, o es importante, o es deseable, o digno de ser deseado. Desde luego, también pueden caber dentro de nuestras explicaciones de esos casos discursos valorativos más «personales»: el agente actúa como actúa porque valora o aprecia la honestidad o la verdad científica. Pero desde la perspectiva del propio agente, en un caso semejante su valoración de la cosa de que se trata es una respuesta razonable por lo menos al valor de la cosa, a su deseabilidad. Y si quien da la explicación del caso comparte el punto de vista del agente sobre el asunto, el nivel de «impersonalidad» de su explicación puede todavía incrementarse: a través de la idea de que el agente hace lo que hace porque reconoce el valor de la honestidad; por ejemplo, porque se percata de que es importante, hasta la idea de que el agente hace lo que hace debido al valor de la honestidad, debido a que la honestidad es importante.

Un agente que actúa así, sólo para enfrentarse después a la duda o incluso al rechazo de otra persona en relación con el supuesto valor, podría intentar, de maneras diversas y contextualmente dependientes, disipar la duda o refutar el rechazo. Quizás no logre ofrecer nada que satisfaga al otro —o que lo satisfaga siquiera a él mismo—; pero aun así puede continuar convencido, modestamente y sin dogmatismos, de que es posible decir *algo* en apoyo de la supuesta deseabilidad del objeto hacia el cual encaminó su acción. No necesita mostrar mayor optimismo acerca de su propia capacidad para apoyar la tesis pertinente de la deseabilidad, del valor «impersonal», que acerca de la buena disposición del otro para apreciar ese apoyo si llegara a ponerse de manifiesto. Sin embargo, ni esa justa precaución dentro de algún contexto coloquial particular, ni siquiera alguna semejante precaución general justa sobre el tipo particular de caracterización de deseabilidad en cuestión, pueden servir para amenazar el carácter distintivo de los deseos del tipo (4) y el correlativo carácter distintivo de las formas en que puede invocarse un discurso valorativo en las explicaciones de los casos en que operan tales deseos. Más aún, es posible que ambos participantes en ese diálogo se encuentren en una y la misma persona; de ser así, y si el diálogo termina con la persistencia del deseo implicado, el agente puede pensar que dicho deseo es por lo menos una respuesta razonable a la deseabilidad o al valor «impersonal» de su objeto. Por lo tanto, cualquier caso de esa índole representa un desafío

formidable incluso para formulaciones cautelosas de identificaciones reduccionistas generales de esta forma: *ser un valor no es más que ser valorado, lo que a su vez no es más que ser deseado.*

3.3 El internalismo del valor

La mayor parte de lo expresado hasta aquí acerca del escepticismo cauteloso de Foot ha seguido la línea direccional, por así decirlo, del querer al valorar; las cosas podrían aclararse más si ahora invertimos esa dirección predominante. La inversión nos pone en contacto con uno de los temas más antiguos de la filosofía, pero nos ayudará tratar de aislar primero ese tema de otros con los que se ha confundido frecuentemente.

Las cuestiones relativas a lo que se ha denominado «la posibilidad psicológica de la conducta moral» (Falk 1947-1948: 493) parecen haber estado enmarañadas en las discusiones éticas de los filósofos por lo menos desde la época en que Sócrates se empeñaba en aplacar las dudas de Glaucón y Adimanto. Cualquiera que sea la evaluación que se haga de la historia antigua de las discusiones filosóficas de dichas cuestiones, creo que está claro que, cuando en el presente siglo surgió lo que se pensó que era poco más o menos el mismo debate, sus términos no se habían afinado mucho. Quizás hubo algún progreso cuando el enfoque se centró en la relación que la motivación guarda con el uso mismo de palabras como «debería», «deber», «obligación». Pero ese avance sólo podía ser pequeño mientras los filósofos se contentaran con un mal uso del concepto de motivación, con la imprecisión en su uso de «uso», con la confianza en una noción supuestamente transparente de «el "debería" moral» y con el empleo de nociones de «deber» y «obligación» supuestamente de igual transparencia. Y aunque en efecto algunas doctrinas insostenibles anteriores sobre la «motivación» fueron en gran medida abandonadas, su lugar en cuanto caso supuestamente indudable de «motivación» fue ocupado generalmente por el ejemplo de la motivación prudencial, que se parafraseó usando nociones mal comprendidas, como «interés» y «ventaja». No obstante, es posible identificar una cuestión razonablemente bien definida que se halla involucrada tanto en aquella historia antigua como en el debate contemporáneo; ésa es la cuestión que me ocupará aquí.

Donald Davidson ha formulado lo que denomina «una forma moderada de internalismo»: «si un agente juzga que sería mejor hacer x que hacer y, entonces quiere hacer x más de lo que quiere hacer y». Acerca de esa doctrina, expresa las siguientes observaciones que, como admite, no son concluyentes:

> Después de todo parece bastante obvio que podemos pensar que *x* es mejor y, no obstante, querer más *y*. [La doctrina] resulta más fácil de cuestionar si se formula de esta manera: si un agente cree que debería (o está obligado a) hacer *x*, entonces quiere hacer *x*; porque, desde luego, con frecuencia no queremos hacer lo que pensamos que deberíamos hacer [...]
>
> Es fácil interpretar [la doctrina] de modo que se vuelva falsa, pero es más difícil creer que no hay una lectura natural que la haga verdadera. Pues a nuestra tendencia a aceptar que a menudo pensamos que deberíamos hacer algo y, sin embargo, no queremos hacerlo, se opone también la tendencia a afirmar que si una persona realmente (sinceramente) cree que debería, entonces su creencia tiene que mostrarse en su comportamiento (y, por lo tanto, desde luego, en su inclinación a actuar o en su deseo). Cuando nos dedicamos a distinguir entre pensar que deberíamos y querer —continúa este razonamiento—, o bien utilizamos la frase «pensar que deberíamos» con el sentido de algo como «pensar que es lo requerido por los patrones habituales de la comunidad», o restringimos el querer a lo que es atractivo desde una perspectiva puramente egoísta o personal. (Davidson 1969: 27 [1995: 44])

Davidson finaliza sus observaciones considerando la posibilidad de que haya alguna otra «palabra o frase con la que, de un modo convincente, podamos sustituir "querer"» en la formulación inicial de su «internalismo moderado».

De hecho, sus observaciones se encaminan en buena parte a evaluar una tesis muy distinta de la doctrina formulada inicialmente: una tesis acerca de la relación entre «pensar que deberíamos» y querer. (Nótese que tal tesis también es distinta de cualquier tesis acerca de la relación entre «pensar que moralmente deberíamos» y querer.) La doctrina formulada inicialmente, el «internalismo moderado», versa de manera explícita sobre la conexión entre *juzgar que es mejor hacer* y *querer más hacer* —es decir, la conexión entre los juicios comparativos sobre la «dignidad» de las acciones y las fuerzas comparativas de los deseos relativos a dichas acciones—. Ahora bien, lo primero, el juicio comparativo, puede surgir en un caso dado del hecho de que el agente adopte cierto valor, del hecho de que valore algo. En tal caso, siguiendo nuestra estipulación terminológica anterior (p. 89), el juicio comparativo implicado expresa una evaluación comparativa por parte del agente. Así pues, en el tipo de caso que me interesa considerar aquí, hay en juego por lo menos cuatro factores posibles:

(1) el hecho de que el agente *valore* algo —la honestidad, por ejemplo, o la claridad—;
(2) el hecho de que el agente tenga un *deseo general* dirigido a casos de aquella cosa, como acciones honestas u opiniones claras;

(3) las *evaluaciones comparativas* que hace el agente de las cosas, en concordancia con su valoración, especialmente las evaluaciones que se expresan en términos de «es mejor hacer», «lo mejor es hacer»; y
(4) los *deseos específicos* del agente relativos a las acciones particulares contempladas.

El problema de lo que llamaré «internalismo del valor» estriba en la explicación de las conexiones, si las hay, entre estos cuatro factores posibles; es decir, abreviando un poco, el problema de dar razón de las conexiones, si existen, entre (1), el hecho de que el agente valore algo, y (4), los posibles deseos específicos del agente en relación con las acciones particulares contempladas. Un extremo «internalismo del valor» sostiene que, salvo por los casos en los cuales un agente, por propia tontería, no logra hacer que coincidan sus deseos y sus creencias, la conexión es en cierto modo necesaria; pero aparte de ese extremo, hay un amplio espectro de opiniones posibles.

En un nivel intuitivo, la plausibilidad de cualquier posición dentro de ese espectro a la que se pueda etiquetar razonablemente de «internalista», requerirá cierta restricción (sin petición de principio) del alcance respecto del tipo de la cosa valorada: una restricción, por decirlo así, a los valores prácticos, una restricción en (1) a las valoraciones prácticas. Pero aunque llegara a formularse e introducirse esa restricción, seguiría siendo digno de notarse que cualquier posición «internalista», como la hemos interpretado aquí, seguirá siendo todavía una posición general acerca de todas las valoraciones prácticas. Me parece muy claro que, así interpretado, el tema del «internalismo del valor» haya sido ciertamente uno de los temas inmersos en el debate, en apariencia más específico, acerca de «la posibilidad psicológica de la conducta moral». Pero también debe destacarse, como paso final hacia la clasificación de los temas iniciales que hemos planteado, que el tema del «internalismo del valor» se ha expuesto aquí en términos (abreviados) de las conexiones de la valoración práctica con deseos específicos, dirigidos a acciones, y por lo tanto, precisamente no en términos de las conexiones de esa valoración con la acción misma. Tal exposición se halla determinada por una serie de consideraciones: (a) mi deseo de evitar introducirme en todas las complicaciones que pueden intervenir entre el deseo y la acción potencial conforme al deseo (pp. 28–34); (b) mi deseo de dejar a un lado el asunto distinto de alguna propiedad «preponderante» que supuestamente se encuentra en ciertos tipos de valoraciones prácticas, y (c) mi adhesión a esa parte de la ortodoxia actual

sobre las razones para actuar que sostiene que el deseo es un componente esencial de cualquier razón semejante para actuar (pp. 68-69).

3.4 Del valorar al querer: deseos potenciales no motivados

Lo que ahora quiero hacer es considerar este tema del «internalismo del valor» a la luz de la categorización de los tipos de deseos que introdujimos en el capítulo precedente y que se empleó en la primera parte de esta discusión sobre el valorar y el querer. Más específicamente, ahora quiero considerar, en relación con cada uno de esos tipos de deseo, si existe alguna posibilidad sustancial de que, en un caso en que podría haber surgido un deseo del tipo de que se trata, puede, sin embargo, ocurrir que el deseo de que se trata no surja y también de que el agente, pese a todo, continúe utilizando el tipo apropiado de discurso valorativo sin poner por ello de manifiesto cierto grado de insinceridad de su parte. En el curso de esta consideración, resultará necesario en ocasiones sustituir la abreviación inicial del problema del «internalismo del valor» por una ampliación, para concentrarnos entonces *o bien* en la cuestión de si existe alguna posibilidad semejante de que un agente continúe valorando algo de ese modo, como en (1), cuando falta el deseo general correspondiente, como en (2), *o bien* en la cuestión de si existe alguna posibilidad de que un agente continúe así con la evaluación comparativa pertinente, como en (3), cuando falta el deseo específico correspondiente, como en (4). Pues en ciertos casos esas cuestiones parecen tener respuestas diferentes.

No es preciso que perdamos tiempo con «el área configurada por potenciales "deseos generadores de razones"», con «el área formada por deseos potenciales de tipo (1)». Pues hemos visto que ningún tipo de discurso valorativo tiene un lugar directo en relación con algún deseo real que pertenezca puramente a este tipo; más aún, hablar de «el área configurada por deseos *potenciales* de este tipo» resulta oscuro, cuando no irrelevante. Pero la subclase restante de los deseos no motivados, (2), exige un examen más detenido.

Con respecto a algunos miembros de esta subclase y en ciertas condiciones —*grosso modo*, cuando el contenido de los deseos en cuestión no es ni trivial ni caprichoso, y cuando el agente en cuestión dedica muchos esfuerzos a alcanzar los objetos de esos deseos—, hemos visto que algunos tipos de discursos valorativos «personales» pueden encontrar un sitio de manera natural: decir que el agente valora o aprecia el objeto de su deseo, decir quizás que le confiere valor a ese objeto, incluso tal vez que ese objeto es importante para él. No obstante, por decir lo menos, resulta difícil

entender cómo un discurso introducido de ese modo puede prestarse luego a crear (por expresarlo así) un *espacio potencial* entre ese valorar y ese querer, un espacio potencial donde fuera posible que el valorar ocurriera en ausencia del desear. Es verdad que un agente con tal deseo —un deseo importuno de justicia social, por ejemplo— posee una caracterización de deseabilidad* (p. 82) del objeto del deseo, una caracterización que no hace referencia lógicamente a la existencia de ese mismo deseo; sin embargo, no hay nada que convierta a esa caracterización en una caracterización de *deseabilidad** excepto la existencia de ese deseo. El asunto, como tentativamente lo sugerimos antes (pp. 89-90), parece comenzar y terminar con que el agente quiera de esta manera ese objeto no trivial, siendo esa manera de querer tal objeto su manera de valorar ese objeto. Y esta última identificación excluye todo espacio, incluso potencial, entre ese valorar y ese querer.

Resultará instructivo, sin embargo, analizar cómo podría efectuarse un intento por introducir tal espacio. Según Judith Baker, alguna vez H.P. Grice consideró el siguiente tipo de solución para el problema distinto de la conexión entre «obligación» y «motivación»: «si Juan piensa que debería realizar una acción *a*, entonces ello exige que o bien Juan quiera *a*, o bien que crea que debería querer *a*» (Baker 1986: 472). La propuesta análoga para el valorar sería ésta: si Juan valora algo, entonces ello exige que o bien Juan desee esa cosa, o bien que valore el tener ese deseo.

Tanto la propuesta que Grice contempló como su análoga para el caso del valorar pasan por alto las complicaciones que surgen de los conflictos, sea entre «obligaciones» o entre valoraciones. De más importancia es que, como la propuesta de Grice, la propuesta análoga tiene que aplicarse consecutivamente: si, en un caso dado, Juan no desea de hecho la cosa valorada, entonces el mismo «análisis» del valorar debe aplicarse al segundo disyunto que esté operando en ese momento. Judith Baker escribe lo siguiente acerca de la propuesta sobre «obligación» y «motivación»:

> En principio ese análisis regresivo no tiene fin. Pero la idea es que en algún punto las personas reales llegarán a detenerse y el juicio de obligación original se concretará en un deseo. Pero no hay garantía de cuál será el nivel en que esto ocurrirá. (*Ibid.*)

La idea comparable que se aplica a la propuesta análoga es que el valorar «se concretará» en un deseo *en algún nivel*: la valoración de Juan exige que quiera la cosa en cuestión, o que quiera quererla, o que quiera querer-quererla, o...

Una vez que se reconoce la posibilidad de deseos de este segundo tipo, el (2), no hay entonces ninguna buena razón para negar la posibilidad de deseos de este tipo de niveles superiores al primero: la posibilidad de deseos de este tipo cuyos contenidos proposicionales se relacionan con la existencia de otros deseos potenciales del agente en cuestión. Así pues, no hay una buena razón para dudar de la coherencia de la explicación propuesta del valorar en los casos que estamos considerando, en los cuales el hecho de que Juan valore algo no se equipara con que lo desee (de este modo), sino que más bien se equipara con el hecho de que *o* lo desea (de este modo), *o* desea (de este modo) desearlo, *o* desea (de este modo) desear desearlo, *o* desea (de este modo)... Esta propuesta afronta, no obstante, una dificultad.

Consideremos la propuesta de que, en un caso dado, el hecho de que el agente valore algo «se concreta» en, o equivale a, el hecho de que quiera (de este modo) quererlo: al hecho de tener un deseo de segundo nivel del tipo (2). Ahora bien, ¿cómo puede ser que el hecho de que el agente tenga ese deseo equivalga a su valorar el objeto del deseo de primer orden inserto en él, y no a su valorar *el tener* el deseo de primer orden? ¿De qué manera puede transmitirse (por así decirlo) la deseabilidad *de este tipo* del tener el deseo de primer orden al objeto de ese deseo potencial de primer orden? Dado que el deseo de segundo nivel es de tipo (2), no hay *ex hypothesi* ninguna razón que pueda aducirse en apoyo de la deseabilidad de su objeto, de la deseabilidad de este tener el deseo potencial de primer orden. El agente solamente experimenta el deseo del segundo nivel, solamente quiere tener el deseo potencial de primer orden. De suerte que, *ex hypothesi*, no hay ninguna razón de esa índole que pudiera servir de alguna manera para transmitir la deseabilidad* en cuestión al objeto del deseo potencial de primer orden. Si lo que el agente simplemente quiere es tener el deseo de primer nivel en cuestión, ese simple deseo se satisface al llegar él a tener ese deseo de primer nivel; y si esto llegara a ocurrir alguna vez, no habría nada que decir, desde el punto de vista establecido por el simple deseo inicial de segundo nivel, con respecto a que el agente busque e intente alcanzar entonces el objeto de ese deseo de primer nivel. Tiene ahora todo lo que, simple y originalmente, quería tener. Desde luego, ahora tiene el deseo de primer nivel en cuestión y, por lo tanto, podría llegar a dedicar un tiempo y un esfuerzo considerables a la consecución de *su* objeto: pero, ¿por qué el agente no ha de ver esto como un efecto colateral, infortunado aunque previsible, de la satisfacción de su deseo simple y original de segundo nivel? Los deseos

de niveles superiores del tipo (2) no pueden constituir la valoración del objeto de un deseo potencial de primer nivel de ese mismo tipo.

Supongamos que un agente tiene un deseo de segundo nivel del tipo (2): sólo quiere querer una cosa. Supongamos, además, que por ello llega de alguna manera a tener el deseo de primer nivel correspondiente. Esta segunda suposición no es tan directa. Dado que el deseo de segundo nivel es de tipo (2), es un simple querer, varios factores que en otras circunstancias podrían haber desempeñado algún papel en la generación del deseo de primer nivel quedan descartados: las caracterizaciones de deseabilidad* completas de los objetos de deseos del tipo (2) no hacen referencia lógicamente a la existencia de deseos que tenga el agente, ni hay posibilidad alguna de que el agente ofrezca alguna razón en apoyo de la tesis de que la caracterización de deseabilidad* completa del objeto de un deseo de este tipo es de tal índole que sirve para captar la deseabilidad de ese objeto. Aun así, una posibilidad es que el agente tome alguna droga que sabe que le producirá el deseo de primer nivel en cuestión.

Pese a las analogías obvias entre el deseo inducido por este medio y los deseos pertenecientes a la primera subclase, el tipo (3), de los deseos motivados, hay también una falta de analogía crucial: en términos generales, ninguna caracterización de deseabilidad completa del objeto del deseo de primer nivel, producido de esta suerte a fin de satisfacer el deseo simple de segundo nivel de tipo (2), necesita hacer referencia lógica a la existencia de ese deseo de segundo nivel. En consecuencia, con respecto a estos casos de producción de deseos de primer nivel, no parece haber otra alternativa, al menos en general, que incluir los productos en la misma subclase de deseos no motivados, el tipo (2), en la cual ocurren los deseos originales de segundo nivel. Podríamos decir que un capricho ha engendrado otro.

¿Qué ocurre con la cuestión de la conexión, en este contexto, entre (3), las evaluaciones comparativas de un agente, y (4), sus deseos específicos en relación con acciones particulares contempladas, como que quiera más hacer *a* que hacer *b*? A primera vista, el asunto parece bastante claro. Omitiendo algunas restricciones de detalle necesarias, (a) se ha estipulado aquí que cualquiera de esas evaluaciones comparativas por parte del agente es consecuente con cierta valoración suya, y (b) se acaba de arribar a la conclusión de que, dentro del área que estamos considerando, cualquiera de esas valoraciones equivale a que el agente tenga el deseo general correspondiente del tipo (2). Por lo tanto, parece que cualquier caso en el cual un agente avala con sinceridad una evaluación comparativa

particular sin tener aún, no obstante, el deseo específico correspondiente, será un caso de irracionalidad por parte de ese agente: pues dado su deseo general y sus creencias acerca del asunto particular de que se trate, parece que un principio de razón práctica exige que debería tener el deseo específico correspondiente. De manera que, aunque sin duda hay un espacio potencial entre la evaluación del agente y el tener el deseo específico correspondiente, ese espacio posee relativamente poco interés; ya que su realización, si bien no basta para la idiotez, es sin embargo suficiente para la irracionalidad. Mientras que si en un caso dado el espacio potencial no se realiza –si el agente llega a tener, en efecto, el deseo específico pertinente–, el deseo en cuestión será un deseo motivado del tipo (3): cualquier caracterización de deseabilidad completa de su objeto, cualquier caracterización que sirva para responder a la pregunta de por qué quiere más hacer *a* que hacer *b*, hará referencia lógicamente al deseo general por parte del agente (o a su valoración, pero esa valoración equivale en estos casos a su deseo).

3.5 Del valorar al querer: deseos potenciales motivados

Podemos comenzar la consideración del área formada por los deseos potenciales del tipo (3) reflexionando sobre la propuesta de que, en esta área, cualquier espacio entre valorar o evaluar, por un lado, y los deseos generales o específicos, por el otro, no es más que una manifestación de irracionalidad por parte del agente en cuestión, no es otra cosa que la manifestación de una transgresión de los principios de la razón práctica cometida por él. Pero por atractiva que sea inicialmente esa propuesta, su desarrollo ulterior requeriría que se llevara a cabo una cantidad de trabajo mucho mayor: en primer lugar, en la consideración de todas las distintas variedades de casos que caen bajo la descripción muy general que se ha dado aquí de los deseos motivados de tipo (3); después, en la formulación de los principios generales de la razón práctica para cada una de las variedades así consideradas; y, finalmente, en la ponderación de la sospecha de que, para que algunos de esos principios sean defendibles, su aplicación necesitará restringirse explícitamente a ciertos tipos de deseos. De suerte que, para no enfrascarme en este momento en los detalles de algún desarrollo de la propuesta mencionada, me quedaré con la idea siguiente: en los casos intuitivamente defendibles de deseos motivados de tipo (3), los deseos así «motivados» son respuestas razonables por parte del agente a los hechos del caso, incluidos los hechos relativos a los otros deseos del agente. Desde este punto de vista, una respuesta racional, una

respuesta exigida por los principios generales de la razón, no es sino *un tipo* de respuesta razonable.

Una caracterización general, pues, de los casos que ejemplifican el espacio potencial en esta área conformada por los deseos potenciales de tipo (3) es la siguiente: dados sus otros deseos y sus creencias acerca de los hechos pertinentes, el agente reconoce el valor conferido por este medio a un objeto, y, sin embargo, aunque no tiene nada en contra de ese objeto, sencillamente no lo desea; admite que el objeto es, por lo menos, digno de ser deseado, que puede desearse razonablemente, pero él, de hecho, no lo desea. Después, como una caracterización de una subclase de los casos posibles seleccionados por esa descripción más general, queda lo siguiente: dados sus otros deseos, sus creencias acerca de los hechos pertinentes y los principios generales de la razón práctica, el agente reconoce el valor conferido por este medio a un objeto, y, sin embargo, aunque no tiene nada en contra de ese objeto, no lo desea; admite que el objeto merece ser deseado, que racionalmente sería deseado, que debería ser deseado, pero él, de hecho, no lo desea. Sobre todo, en cualquier caso dado en que un agente encuentre que se realiza el espacio potencial general entre el valorar y el desear de tipo (3), podría entonces pasar a preguntarse cómo se puede eliminar en este caso el espacio en cuestión, cómo se puede provocar en este caso el deseo faltante. Podría considerar si existe algún rasgo cambiante de su vida mental o personal que esté operando para obstruir el deseo potencial en cuestión; o podría considerar el intento de concentrar toda su atención en el otro deseo «motivante» que de hecho tiene, el intento de concentrarse en su fuerza o en el atractivo o la deseabilidad de su objeto; o podría considerar el intento de concentrar toda su atención en las exigencias de los principios de la razón práctica que estén en juego; o, de nuevo, podría considerar recurrir a alguna droga apropiada. Y si tal consideración deriva en una acción, y si esta acción tiene entonces como resultado el deseo en cuestión, el agente verá ese resultado, en grados variables y de formas variables, como una respuesta o razonable o racional de su parte a las circunstancias del caso. Pero, al igual que antes, esa manera de contemplar los deseos del tipo general (3) no requiere que un proceso de reflexión semejante haya precedido a su existencia.

Al considerar casos en los cuales se realiza este espacio potencial general entre el valorar y el desear de tipo (3), quizás podría parecer tentador tratar de invocar de nuevo el análogo para el valorar de la propuesta contemplada por Grice: esto es, tratar de invocar la idea de que si bien

el agente en cuestión no necesita tener ningún deseo por el objeto cuyo valor reconoce, en cualquiera de estos casos el agente debe o querer tener ese deseo, o querer querer tenerlo, o... Pero cualquiera que sea el valor final de tal invocación, debe destacarse que por lo menos una de sus posibles motivaciones no sería una manifestación de nada mejor que una confusión. La motivación es la de hacerle continuada justicia al escepticismo cauteloso de Philippa Foot acerca de la existencia de una explicación de «lo que queremos decir cuando hablamos de que *alguna cosa* tiene valor, o incluso de que la valoramos, en tanto opuesto a quererla o a estar dispuesto a tomarse la molestia de conseguirla». Aquí la cuestión no es simplemente que la invocación sugerida no capte con exactitud el pensamiento escéptico al que se alude ahí; es también que en la medida en que esa invocación se aproxima a ese pensamiento, también se aproxima a él la explicación que hasta ahora hemos dado. Pues según esta explicación, cualquier juicio pertinente de deseabilidad o de valor que haga un agente cuando carece del deseo específico correspondiente, se funda de todos modos en algún otro deseo que realmente tiene: la caracterización de deseabilidad completa del objeto de ese juicio, del objeto del deseo meramente potencial, hará referencia lógicamente a la existencia del otro deseo que el agente tiene de hecho.

Fue en el área determinada por los deseos motivados de la clase restante, el tipo (4), donde encontramos que cabían con más naturalidad las formas más «impersonales» de discurso valorativo. En cualquier caso de un deseo de este tipo, la caracterización de deseabilidad completa del objeto de ese deseo no necesita hacer referencia lógica a la existencia de ningún deseo que tenga el agente en cuestión; pero, desde un punto de vista más positivo, para cualquier deseo de este tipo existe la posibilidad de que el agente, dentro de cierto contexto específico, aduzca alguna razón que apoye la afirmación de que la caracterización de deseabilidad completa de su objeto es tal que sirve para captar la deseabilidad de su objeto. Parece difícil negar que esa característica definitoria positiva de los deseos de tipo (4) está de alguna manera conectada con las formas distintivas en que el discurso valorativo se introduce de manera natural en la explicación de casos reales de acción conforme a deseos de este tipo.

Cuando aparece un deseo cualquiera de este tipo (4), se ponen potencialmente en juego por lo menos los tres factores siguientes (pp. 68–73):

(a) el deseo general del agente —de honestidad, por ejemplo, o de verdad científica—;

(b) la caracterización de deseabilidad general que el agente tiene o acepta del objeto de ese deseo —la caracterización general, y generalmente «impersonal», de ese objeto, en la que se hace una invocación a valores—; y

(c) las varias razones que el agente puede aducir en contextos diversos para apoyar la afirmación de que la caracterización de deseabilidad completa del objeto de deseo es tal que sirve para captar la deseabilidad de ese objeto, en apoyo de esa invocación de un discurso valorativo «impersonal».

Ahora bien, ya antes nos vimos llevados a aceptar el carácter sustantivo de la distinción exclusiva, enteramente general, entre *desear* un objeto y *tener una caracterización de deseabilidad* de él (pp. 62-63); así que tenemos que aceptar el carácter sustantivo de esa distinción general en la aplicación a los casos del tipo (4), el carácter sustantivo de la distinción entre (a) y (b) en tales casos. Esa aceptación exige otra: la del carácter sustantivo de la cuestión, relativa al área conformada por los deseos potenciales del tipo (4), de si hay casos posibles en los que ocurran (b) y (c) en ausencia de (a) —la cuestión, relativa a esa área, de la posibilidad de que haya casos de afirmaciones sinceras basadas en razones en cuanto a valores, y de reconocimiento de valores sin deseo—. Más aún, el nivel más elevado de generalidad en la descripción de casos en que en efecto ocurren deseos de este tipo (4) parece sugerir que aquí hay posibilidades reales. Ese nivel de descripción discurre como sigue. Alguna razón, como en (c), suscita razonablemente, produce como respuesta razonable, la aceptación de una caracterización general de deseabilidad como en (b); y esa aceptación, a su vez, suscita razonablemente, da lugar razonablemente, al deseo correspondiente como en (a) (pp. 71-73). A través del reconocimiento de ese proceso, podemos comprender de un modo preciso y natural por qué el agente tiene el deseo que de hecho tiene. Pero, si esto es así, parece entonces que nada tiene de imposible un proceso de «generación» más o menos similar que se detenga con la aceptación de una caracterización general de deseabilidad por parte del agente, sin que esa aceptación dé lugar, de hecho, al deseo correspondiente.

Acerca de ese *parecer*, debemos señalar primero dos cuestiones. En primer lugar, en esta área habrá a lo sumo un ámbito muy limitado para reemplazar la noción general de una respuesta razonable por la más específica de una respuesta racional, por la de una respuesta exigida por los principios genérales de la razón. En lo tocante a la relación o las relaciones potenciales entre casos de (c) y de (b), la cuestión es obvia: el carácter,

dependiente del contexto, de cualquier razón del tipo (c) representa un obstáculo tremendo para el intento de considerar una relación semejante en términos de principios generales de la razón. El asunto de la relación o las relaciones potenciales entre casos de (b) y de (a) es menos claro, así que nos limitaremos por el momento a señalar la pasmosa historia de fracasos, de puras baladronadas, en los intentos para mostrar que existe algún tipo de irracionalidad en todos los casos aparentes en que (b) ocurre en ausencia de (a). La segunda cuestión que debe hacerse notar se relaciona con el asunto de la explicación de la diferencia entre casos que, en términos generales, son similares en otros respectos, y en los que la diferencia está en si la aceptación de la caracterización de deseabilidad general da o no lugar de hecho al deseo correspondiente. Sin duda, en muchos de esos casos, la explicación que ha de darse hará referencia a otros rasgos mentales y personales de los agentes en cuestión; y sin duda, en muchos de esos casos las formas precisas en que se hace tal referencia revelarán que lo que está en juego son otras valoraciones variables por parte de los agentes en cuestión. Pero no sé de ningún buen argumento que sirva para descartar la posibilidad de casos en que no haya explicaciones de las diferencias de los estados conativos y de las diferencias de los deseos de los agentes que se expresen en términos de su vida mental y personal.

Una inquietud más apremiante acerca de aquel *parecer* surge de la siguiente manera. De acuerdo con lo dicho aquí hasta el momento, los casos en que no se presenta el deseo general, como en (a), pueden separarse en dos grupos toscamente delimitados: aquellos en los que la apreciación, por parte del agente, de una razón, como en (c), da lugar a su aceptación de una caracterización de deseabilidad general, a algún presunto reconocimiento de algún valor, generalmente «impersonal», y aquellos en que esto no sucede. Pero, ¿cuál es exactamente el contenido de la distinción así trazada? ¿Cuál es el contenido de ese discurso sobre la aceptación de una caracterización de deseabilidad general, de ese discurso valorativo, en ausencia del deseo general correspondiente? ¿No es el hecho de que el agente tenga el deseo correspondiente el criterio que indica que acepta sinceramente una caracterización de deseabilidad general y la sinceridad de su discurso sobre los valores?

Pero el escepticismo general que plantea esa inquietud es vulnerable en la naturaleza del caso a la amenaza de la posibilidad pasada por alto. Aún más, no hay una buena razón para suponer que tenga que haber en todos esos casos *un* criterio que sirva para marcar la distinción que estamos discutiendo. Así, por ejemplo, mucho dependerá del nivel de

«complejidad estructural» de la caracterización de deseabilidad general en cuestión en el caso de que se trate. En un caso en el cual haya un nivel apreciable de dicha complejidad, la sinceridad de la aceptación del agente de esa caracterización podría ponerse de manifiesto a través de su aceptación de otras caracterizaciones de deseabilidad relacionadas, en cuyo caso dicha aceptación en efecto daría lugar, razonablemente, a los deseos correspondientes de su parte.

Al lector se le habrá ocurrido otra posibilidad: la de volver a invocar algún análogo para el valor de la propuesta distinta contemplada por Grice. Se podría sugerir, por ejemplo, que la aceptación sincera de una caracterización de deseabilidad general, por parte de un agente, tiene que «concretarse» en términos de algún deseo real del agente *en cierto nivel*: el agente tiene que tener el deseo general correspondiente, o querer tenerlo, o querer querer tenerlo, o...

Estamos considerando los deseos potenciales y reales del tipo (4) y las formas distintivas en que el discurso sobre deseabilidad y sobre valor puede entrar de manera natural en las explicaciones de casos del área conformada por los deseos potenciales de ese tipo. Consideremos, pues, un caso dado de esta área en el que un agente juzga que algo posee valor: piensa, por ejemplo, que la honestidad tiene valor, que es «impersonalmente» deseable, y tiene potencialmente varias razones que podrían aducirse, dentro de diversos contextos, en apoyo de esa idea. Si esto es así, entonces también pensará, digamos, que *actuar honestamente* posee valor. Pero sin duda también ocurrirá que el agente en cuestión pensará que *desear la honestidad* posee valor (al igual que desear acciones honestas). Y sin duda también ocurrirá que pensará que *desear desear la honestidad* tiene valor. Y así sucesivamente. La manera precisa en que ha de entenderse ese proceso de «transmisión» bien podría depender de cualesquiera puntos de vista generales sobre el valor que el agente en cuestión sostenga, o del carácter del valor particular implicado, o de ambas cosas. Así, por ejemplo, una posibilidad en cualquier caso particular dado es que (cada eslabón de) la cadena de deseos se considera de valor sólo porque se estima que la aparición de (cada eslabón de) esa cadena constituye un medio necesario para la ejemplificación en el mundo real del valor de que se trata; una posibilidad distinta es que las razones potenciales que podrían aducirse dentro de diversos contextos en apoyo de la pretendida «categoría de valor» de los valores en cuestión, sirven también para apoyar la idea de que (cada eslabón de) la cadena de deseos posee cierto valor «independiente»; mientras que otra posibilidad más es que esas razones

sirven para revelar que la «categoría de valor» del valor en cuestión y de sus ejemplificaciones en (digamos) el actuar se deriva, de cierta forma, del valor de (cada eslabón de) la cadena de deseos. Sin embargo, repito, no parece haber nada incoherente en la idea de que cuál sea la posibilidad razonablemente favorecida, o de que cuál sea la combinación de posibilidades razonablemente favorecida, podría depender del carácter del valor particular en discusión. No obstante, sigue en pie la conclusión de que, dentro de esta área conformada por deseos potenciales del tipo (4), la idea de que algo posee valor va acompañada por la idea de que (cada eslabón de) la cadena de deseos correspondientes posee valor. Y entonces, al reflexionar sobre esa cadena potencialmente infinita de pensamientos valorativos, puede surgir la propuesta de que, para la «gente real», algún miembro de esa cadena de pensamientos valorativos será «concretado» en términos del deseo correspondiente en el nivel contiguo superior de la cadena. Más aún, podría parecer que esa propuesta, formulada ahora dentro del contexto de las reflexiones inmediatamente precedentes, nos ha aproximado a la apreciación de algún *argumento* en favor de la verdad del análogo para el caso del valor de la propuesta contemplada por Grice.

Por desgracia, las dificultades persisten independientemente de la oscura categoría de la tesis sobre la «gente real». Una de ellas está relacionada con la idea inicial de que alguien que piense que la *honestidad* posee valor también pensará que *desear la honestidad* posee valor. ¿Por qué no bastaría, por lo menos en ciertos casos, con que el agente pensara que *valorar la honestidad* tiene valor (y así sucesivamente para (cada eslabón de) la cadena de valoraciones)? La amenaza de petición de principio resulta clara. En casos en los que el desear-*o*-valorar se estima de valor porque es un medio necesario para la ejemplificación en el mundo real del valor en cuestión, esta dificultad puede evitarse: pues dada (a) la aceptación de una parte de la ortodoxia actual acerca de las razones para la acción en cuanto al papel insuprimible de los deseos dentro de tales razones, junto con (b) algún principio verosímil de la razón práctica en términos de medios y fines, se sigue que en tales casos un agente que se conforma con la cadena de (meras) valoraciones se manifiesta por ello como irracional. Pero carecemos todavía de razones para generalizar a partir de tales casos, para suponer que todos los pensamientos sobre el valor del desear o del valorar deban explicarse en función de esos casos. Y esto se vincula con una segunda dificultad relativa a la propuesta de que, para la «gente real», algún miembro de la cadena pertinente de pensamientos valorativos «se concretará» en términos del deseo correspondiente en el nivel contiguo superior de esa cadena. Indudablemente, para la «gente

real» esa cadena no puede continuar para siempre; pero, ¿por qué no podría simplemente detenerse sin esa «concreción»? Al igual que antes, en casos en los que el desear-*o*-valorar se estima de valor porque constituye un medio necesario para la ejemplificación en el mundo real de los valores en cuestión, la ausencia de «concreción» en un deseo pondrá de manifiesto cierto grado de irracionalidad por parte del agente en cuestión; pero, como antes, carecemos todavía de razones para generalizar a partir de tales casos.

La conclusión provisional es, pues, que si bien algún análogo de la propuesta contemplada por Grice para el caso del valor puede servir como un criterio de la aceptación sincera por parte de un agente de un discurso valorativo dentro de esta área, no hay razón para creer que éste o cualquier otro sirva como *el* criterio de tal aceptación dentro de esta área. Cualesquiera que sean las conexiones dentro del área conformada por los deseos potenciales del tipo (4) entre el valorar y el querer, se resisten a ser encapsuladas en una fórmula identificativa neta del tipo que se esconde tras el cauteloso escepticismo de Foot.

3.6 El sentido de la subjetividad, de la dignidad, del mérito

La gente afirma en ocasiones que los asuntos de valores son «subjetivos», y quizás lo dicen más a menudo cuando se trata de asuntos relativos a valores morales que de otros. También los filósofos dicen frecuentemente esas cosas. No necesito prestar mucha atención a esas voces, provengan de boca de gente normal o de los filósofos. No obstante, creo que, en el intento por alcanzar los objetivos declarados de este ensayo, podría ser de ayuda considerar algunas de las cosas que podrían querer estar diciendo esas voces. Al hacerlo, dejaré de lado deliberadamente la cuestión relativa a «el uso correcto» de la palabra «subjetivo».

En los casos de valoraciones dentro del área conformada por los deseos potenciales del tipo (2), hay un significado que podría vincularse de inmediato con la afirmación de que esas valoraciones son subjetivas: el significado encerrado en la idea de que, en el caso de semejante valoración, no hay nada que decir en apoyo de esa valoración, junto con la idea de que, si se produce una diferencia de opinión acerca de esa valoración, no hay nada que pueda decirse a fin de buscar una resolución razonable de la diferencia. Las dos ideas allí implicadas son correctas. Además, hay otras verdades sobre las valoraciones de este tipo que el defensor de la tesis de la subjetividad podría querer estar diciendo. Una de ellas es que, en cualquier caso particular de este tipo, el hecho de que el agente valore

algo consiste justamente en que lo desea (de esta manera); otra es que, en casos de este tipo, no hay un sitio natural para el discurso valorativo «impersonal». Con un poco más de ambición, por medio de la tesis de la subjetividad podría estar expresándose la idea verdadera de que en estos casos no hay nada que decir en favor del pensamiento de que la experiencia del valorar puede contar como un caso de estar en presencia de alguna propiedad de valor o rasgo de valor del mundo que se encuentra allí de modo independiente —que se encuentra allí para ser experimentado independientemente de cualquier experiencia particular de valoración—. Y también podría estar expresándose la verdad relacionada de que en los casos de esta área no hay nada que decir en favor del pensamiento de que los deseos del agente pueden ser respuestas razonables a propiedades o rasgos de valor independientes que se encuentran en el mundo.

Hume escribió que «cuando reputáis una acción o un carácter como viciosos, no queréis decir otra cosa sino que, dada la constitución de vuestra naturaleza, experimentáis una sensación o sentimiento de censura al contemplarlos» (1888: 469 [1977: 689]). Y continúa: «Por consiguiente, el vicio y la virtud pueden compararse con los sonidos, los colores, el calor y el frío, que, según la filosofía moderna, no son cualidades en los objetos, sino percepciones en la mente.» Aquí no es de nuestro interés todavía el asunto de las valoraciones morales y los juicios morales; pero si interpretamos las observaciones de Hume como observaciones sobre casos de valoración del área conformada por los deseos potenciales del tipo (2), y al mismo tiempo hacemos a un lado su referencia al «significado», las opiniones expresadas así se aproximan mucho a las verdades subjetivistas que acabamos de reconocer. Esas valoraciones no son casos en que se nos presenten «cualidades en los objetos», ni los deseos que están en juego son respuestas razonables a cualesquiera cualidades de esa índole. En realidad, con esas valoraciones y esos deseos puede ser que necesitemos escuchar la advertencia de Hume contra el riesgo de esparcirlos sobre el mundo que existe ahí de un modo independiente, el peligro de objetivizar nuestras valoraciones y deseos al tomarlos, precisamente, por experiencias de cualidades de valor fuera de la mente y por respuestas razonables a ellas. Pero sería dar un paso equivocado concluir que lo que de esta forma se valora está ello mismo «en la mente», es la misma experiencia o «percepción» de la valoración (véase McGinn 1983: 130). Ninguna valoración de este tipo cuenta como un caso en que se experimenta una cualidad de valor o rasgo de valor que ha de hallarse en los objetos fuera de la mente; no obstante, lo valorado, el objeto de

nuestras predicaciones de valor, puede ser, sin embargo, algo fuera de la mente, puede ser, sin embargo, cierta «acción o carácter» de otro.

Las interpretaciones posibles de la tesis de que las valoraciones son subjetivas que la vuelven verdadera en relación con las valoraciones dentro del área conformada por los deseos potenciales del tipo (2), no la vuelven verdadera en relación con las valoraciones características del área conformada por los deseos «motivados» potenciales del tipo (3). En casos de este tipo hay algo que se puede decir en apoyo de una valoración semejante y con el propósito de producir una resolución razonable de cualquier diferencia relativa a tal valoración: por medio de la referencia a la existencia del otro deseo en juego, el deseo «motivante», junto con la referencia a los presuntos hechos relevantes en cuestión, como las consideraciones instrumentales o constitutivas pertinentes. De esta forma se puede producir una caracterización de deseabilidad de la cosa valorada en casos de este tipo que no haga referencia ni al hecho de que el agente realmente la valora ni al de que realmente la desea. Esa caracterización pretende revelar por qué esa cosa es, por lo menos, digna de ser valorada; ciertamente, tal caracterización podría pretender incluso revelar por qué esa cosa debería ser valorada, por qué merece ser valorada. En casos de este tipo, hay también un espacio para distinguir entre la valoración del agente de la cosa en cuestión y su deseo real de ella, y, a la vez, un sitio natural para el discurso valorativo de cierto nivel de «impersonalidad». Finalmente, en tales casos pueden de vez en cuando surgir razonablemente las siguientes ideas: la experiencia del valorar cuenta como un caso en que se nos presenta algún valor —propiedad de valor o rasgo de valor del mundo que está ahí de manera independiente—, el cual está ahí para ser experimentado independientemente de cualquier experiencia particular de valoración; y el deseo del agente por la cosa en cuestión es una respuesta razonable a propiedades de valor o rasgos de valor del mundo independientes —independientes, digo, de la experiencia particular del agente de querer la cosa en cuestión y de valorarla—.

Sin embargo, esta huida de la subjetividad sólo nos ha hecho avanzar un tramo corto, como lo pone de manifiesto la consideración de casos en que el otro deseo en juego, el «motivante», es del tipo (2). Pues, en términos generales, en cualquier caso semejante tanto la fuerza del presunto apoyo de la valoración del agente como la plausibilidad de la esperanza de una resolución razonable de cualquier diferencia en tal valoración, redundan en la cuestión fáctica —de meros hechos, no razonables— de si otros comparten este otro deseo «motivante» de tipo (2). Se ha aceptado

aquí que tal cuestión es subjetiva en más de un sentido: en particular, se ha aceptado tácitamente que si ese compartir no ocurre de hecho, no hay nada que pueda decirse o hacerse con miras a producir una modificación razonable en ese asunto. De suerte que, en términos generales, si ese compartir no ocurre de hecho, tanto la fuerza del presunto apoyo de la valoración del agente como la plausibilidad de la esperanza de una resolución razonable de cualquier diferencia en tal valoración parecen evaporarse. (Esto es así sólo *en términos generales*: pues en ciertos casos existe una posibilidad clara de una crítica razonable *del deseo motivado* del tipo (3), incluso por parte de alguien que no comparta el deseo «motivante» que está en juego.)

Nada de lo anterior exige una retractación de las otras afirmaciones que se acaban de introducir en relación con las valoraciones dentro del área conformada por los deseos potenciales del tipo (3), ni siquiera cuando el otro deseo en juego, el «motivante», es de tipo (2). Pero en los casos en que este otro deseo es de ese tipo o del tipo (1), las demás afirmaciones introducidas se antojan relativamente anodinas. Aunque, por ejemplo, pueda permitirse en tales casos cierto discurso sobre valores «independientes», el discurso así permitido tiene que interpretarse de un modo sumamente cauteloso y modesto: tiene que interpretarse meramente en términos de la independencia de la experiencia particular del agente particular de querer la cosa de que se trate y de valorarla. Así interpretado, ese discurso resulta del todo compatible con la característica insistencia del subjetivista en que las valoraciones y deseos humanos reales constituyen la fuente de todo valor: simplemente sucede que en estos casos esa fuente, por así decirlo, se esparce sobre el mundo más ampliamente que en otros casos más simples.

En consecuencia, adquiere mucho mayor urgencia considerar los presuntos casos en que el otro deseo, el «motivante», es él mismo un deseo «motivado» del tipo (4); esto a su vez vuelve mucho más urgente considerar la tesis de la subjetividad en relación con las valoraciones dentro del área conformada por los deseos potenciales de ese tipo. En los casos de esta área, hay de nuevo cosas que pueden decirse en apoyo de cualquier valoración semejante y con el propósito de producir una resolución razonable de cualquier diferencia en esa valoración: es decir, pueden aducirse razones que apoyen la idea de que la caracterización de deseabilidad expresada es en verdad capaz de captar la deseabilidad de la cosa en cuestión, es en verdad una caracterización que sirve para revelar que la cosa en cuestión es, por lo menos, digna de ser valorada. Pero

en contraste con los casos inmediatamente precedentes dentro del área conformada por los deseos potenciales del tipo (3), en los casos presentes las cosas que pueden decirse no harán ahora referencia a la existencia de algún otro deseo que el agente tenga: serán más bien intentos por llevar y concentrar la atención sobre algunas «cualidades (de valor) en los objetos» fuera de la mente, sobre algunos rasgos de valor de la realidad, independientes de los deseos y valoraciones reales del agente. La distinción en estos casos entre valoración y deseo; el sitio natural que en estos casos ocupa el discurso valorativo completamente «impersonal»; la idea de que en estos casos valorar puede considerarse un ejemplo en el que se nos presenta alguna propiedad de valor o «cualidad en los objetos» que de todas formas está ahí y que es por lo menos digna de ser valorada independientemente de los deseos y las valoraciones humanas reales; la idea de que en estos casos el hecho de que un agente desee la cosa en cuestión puede ser por lo menos una respuesta razonable al valor que esa cosa tiene de manera independiente: todo esto sirve ahora para marcar una distancia considerable dentro de la huida de la subjetividad. Dicho de la manera más concisa posible: nuestra comprensión de tales casos resulta francamente incompatible con la propuesta, entendida descriptivamente, de que las valoraciones y los deseos humanos reales son la fuente de todo valor. Además surgiría una incompatibilidad estructuralmente similar con cualquier propuesta subjetivista semejante si hubiera que considerar los casos de valoración dentro del área conformada por los deseos potenciales del tipo (3) en los que el otro deseo en juego, el «motivante», es de tipo (4).

También es crucial reconocer que quedan varios significados posibles que podrían vincularse con la tesis de la subjetividad de acuerdo con los cuales esa tesis sigue siendo compatible con la explicación expuesta aquí de los casos de valoraciones dentro del área conformada por los deseos potenciales de este tipo final, el (4). En primer lugar, la tesis del subjetivista podría consistir en que tanto el asunto de la capacidad para las valoraciones de este tipo, como el de determinar precisamente cuáles de estas valoraciones se hacen de hecho, dependerán de la naturaleza del «sujeto», de la naturaleza de la persona implicada: de su constitución fisiológica y psicológica, de sus necesidades, preocupaciones e intereses. Y en tanto que esa tesis no se formule en términos concebidos para eliminar el área conformada por los deseos potenciales del tipo (4) en favor de la conformada por los deseos potenciales del tipo (3), no hay necesidad de negarla aquí. Ello se relaciona con una segunda verdad posible oculta tras

la tesis de la subjetividad. Hemos sostenido aquí que en los casos de valoraciones que están dentro del área conformada por los deseos potenciales del tipo (4), hay cosas que se pueden decir con el propósito de producir una resolución razonable de cualquier diferencia en tal valoración: pero con ello no pretendíamos insinuar un gran optimismo ni respecto de la frecuencia de la realización de ese propósito ni respecto de que su realización siempre resultara verdaderamente posible. Existen rasgos de quien aduce razones y rasgos de quien recibe razones que pueden combinarse fácilmente para malograr la realización de ese propósito, y los rasgos de uno o del otro podrían bastar en un caso dado para volver irrealizable ese propósito. Más aún, nada de lo aquí expuesto sobre los casos de esas valoraciones tiene la intención de autorizar la arrogancia, la complacencia o la certidumbre generalizada por parte de un sujeto acerca de sus valoraciones reales de este tipo en un momento dado. La verdad es exactamente lo opuesto: una vez que se toma en serio la idea de que tales valoraciones pretenden ser respuestas cognoscitivas a las «cualidades en los objetos», cualidades que se encuentran ahí de manera independiente, está claro que la precaución estará entonces a la orden del día. Nuestra sensibilidad hacia el modo de ser del mundo siempre puede mejorarse, en éste como en cualquier otro caso. La huida de la subjetividad que aquí se ha emprendido no conduce a desembarcar en la aceptación de un dogma (véanse Platts 1979: 247; 1980b: 70–71).

Apenas valdrá la pena, ahora, desarrollar esas cuestiones; pero aún falta considerar otro asunto relacionado con ciertas preocupaciones posibles de quienes sostienen la tesis de la subjetividad. Esto atañe al contenido, en los casos de valoraciones dentro del área conformada por los deseos potenciales del tipo (4), de la idea de que el objeto de cualquier valoración semejante se estima digno de ser deseado. Quizás la cuestión principal que debe destacarse es la debilidad relativa (en cierto sentido) de esa idea. Sostener que algo es digno de ser valorado y digno de ser deseado no equivale todavía a sostener que debería ser valorado y debería ser deseado: es decir, no es equivalente de sostener que merece ser valorado y ser deseado. (Creo que el elemento de estipulación lingüística en esa manera de plantear el asunto es ligero y, de todos modos, inocuo.) La tesis es más bien que la cosa en cuestión es un objeto digno de ese valorar y ese desear: pero hay demasiadas cosas que son objetos dignos de tal valorar y tal desear. Pensemos, por ejemplo, en la búsqueda de la verdad histórica acerca del Renacimiento del siglo XII: puede sostenerse al mismo tiempo, muy razonablemente, que quien se dedica a la búsqueda de esa verdad se dedica a un fin digno, y también que, por así decirlo, a alguien

que vive en una bendita ignorancia hasta de la existencia de ese tema de estudio no se lo puede criticar de forma razonable. Más todavía, aun cuando una persona reconozca en realidad el valor de la búsqueda de la verdad respecto de ese tema de estudio, sin que posea aún ningún deseo de dedicarse él mismo a ella, está muy lejos de estar claro que haya base alguna para criticarlo razonablemente.

Podría resultar útil explicar en mayor detalle las diversas formas en que este último tipo de casos, caracterizado en líneas generales, podría desarrollarse más. Una posibilidad estriba en que la persona implicada o bien valore la dedicación de otros a la búsqueda en cuestión, o bien desee esa dedicación. Otra posibilidad consiste en que dicha persona o bien valore la «ausencia de obstáculos» en la actividad de otros que están dedicados de hecho a esa búsqueda, o bien desee tal «ausencia de obstáculos». Y otra posibilidad es que la persona implicada o bien valore dar asistencia a aquellos otros que están de hecho dedicados a esa búsqueda, o bien desee dar tal asistencia. Pero todos esos casos, según se han descrito, son muy distintos de aquel en que se invoca, de un modo defendible, la idea de que la cosa originalmente valorada merece ser valorada y merece ser deseada, y *debido* a ello tiene el ser valorada y el ser deseada: cualquier caso en que la cosa originalmente valorada debería ser valorada y debería ser deseada.

No resulta problemático invocar por lo menos algún pensamiento de este tipo para la descripción de ciertos casos reales. Piénsese, por ejemplo, en el área conformada por deseos potenciales del tipo (3) y en los casos, dentro de ella, en que el deseo «motivante» pertenece al tipo (1) o al tipo (2): puede invocarse naturalmente, por lo menos a veces, una afirmación sobre lo que un agente debería valorar y debería desear. Pero, desde luego, tal tesis se encuentra condicionada, o es relativa, a la ocurrencia del otro deseo, el «motivante». Mucho más tendenciosas serían las afirmaciones sin relativización alguna en el sentido de que toda persona debería reconocer algún valor de este tipo, de que toda persona debería valorar (de esta manera) la cosa de que se trata, y en el sentido de que toda persona debería desear la cosa en cuestión: de suerte que tanto la ignorancia como la «falta de voluntad» resultarían siempre objetos de crítica potencialmente razonables. Nada de lo dicho hasta aquí tiene la intención de apoyar o refutar la tesis de que tales afirmaciones pueden en ocasiones enunciarse con verdad. De manera que si la preocupación subyacente tras la tesis de la subjetividad se refiere a la negación de aquella tesis, está discusión no ha logrado hasta el momento hacer frente a dicha preocupación. La intención de lo que se acaba de decir ha sido más

bien tratar de elucidar la debilidad comparativa de las varias afirmaciones que aquí se han enunciado, empleando la noción de dignidad, no la de mérito.

3.7 ¿Se trasciende lo subjetivo sensible?

Una cuestión final de posible interés para el subjetivista se relaciona con la aparición de lo que podría denominarse «trascendentalismo» dentro de la explicación dada hasta aquí de las valoraciones del área conformada por los deseos potenciales del tipo (4). En vista de lo que se ha dicho en defensa de la idea de que un caso de una valoración semejante puede contar como respuesta cognoscitiva a una propiedad de valor o a un rasgo de valor que se halla en forma independiente en los objetos fuera de la mente, podría suscitarse la idea de que esta explicación sostiene que la existencia de esas propiedades de valor o rasgos de valor —o, en una abreviación aparentemente inocua, la existencia de valores— es por completo independiente de las valoraciones y los deseos humanos. Y esto podría juzgarse como algo bastante difícil de digerir.

Al igual que tantas otras cosas en esta área, quizás abordemos mejor la cuestión si retornamos a Hume. Éste sostenía, como hemos visto, que el vicio y la virtud son por lo menos comparables provechosamente con lo que se ha llegado a conocer como cualidades secundarias, como los colores, que «según la filosofía moderna, no son cualidades en los objetos, sino percepciones en la mente». Esa comparación, junto con la presunta concepción de las cualidades secundarias, ha sido examinada instructivamente en varios escritos filosóficos recientes, por algunos de los *más modernos* filósofos, de suerte que, con su ayuda, puedo abreviar aquí mi exposición. Pero tengo que enfatizar desde el principio que aquí todavía no me preocupa, sin embargo, el tema de los valores morales y los juicios morales: mi tema sigue siendo el muy general de las valoraciones dentro del área conformada por los deseos potenciales del tipo (4). ¿Qué podremos hacer, pues, con las observaciones de Hume si de manera deliberada las interpretamos erróneamente como si estuvieran dirigidas a este tema?

La primera cuestión que debe destacarse es que en su discusión sobre las cualidades secundarias como los colores, Hume, al igual que muchos escritores antes y después que él, desarrolla juntos, de un modo ilegítimo, dos ideas muy diferentes que podrían expresarse diciendo que esas cualidades son «subjetivas». Una es la idea de que, en palabras de John McDowell, una cualidad secundaria «es una propiedad cuya atribución a un objeto no se comprende adecuadamente excepto como verdadera,

si es verdadera, en virtud de la disposición del objeto a presentar cierto tipo de apariencia perceptiva» (McDowell 1982: 111). La otra consiste en que tales cualidades «no son cualidades en los objetos», de suerte que la experiencia de tales cualidades no puede ser una conciencia perceptiva de propiedades que genuinamente posean los objetos. Una vez que se distinguen ambas ideas, no parece haber ninguna buena razón para pensar que la aceptación del primero exija la aceptación del segundo: ninguna buena razón, esto es, para creer que haya alguna incoherencia en la idea de una propiedad que los objetos posean genuinamente y cuya atribución a los objetos, no obstante, haya de entenderse en términos de que esos objetos tengan una disposición a producir ciertas «percepciones en la mente». La ambigüedad de la pregunta «¿Hay cualidades secundarias fuera de la mente?» hace visible la cuestión con nitidez. La pregunta puede responderse afirmativamente en el sentido de sostener que tales cualidades existen *fuera de* la mente, que son en verdad propiedades que los *objetos* poseen genuinamente y, no obstante, al mismo tiempo puede responderse de manera negativa en el sentido de sostener que cualquier elucidación verídica acerca de nuestra comprensión de las atribuciones de esas cualidades tiene que *hacer referencia a* los efectos potenciales producidos en las mentes, producidas en los *sujetos*.

Nada de lo anterior pretende negar la posibilidad de que haya cualidades que sean subjetivas en los dos sentidos distinguidos; pero, no obstante, supongamos, y hay buenas razones para suponerlo, que las cualidades secundarias como los colores son subjetivas sólo en el primero de esos sentidos, aceptando con ello que se trata de propiedades que los objetos poseen genuinamente. Y observemos ahora, dentro del contexto de esa suposición, que la propuesta de Hume es meramente que hay una comparación útil entre las cualidades secundarias y las experiencias perceptivas pertinentes, por un lado, y los valores y las valoraciones que caen dentro del área conformada por los deseos potenciales del tipo (4), por el otro. Es decir, sería a lo sumo un magro triunfo sobre la propuesta de Hume conformarse únicamente con señalar que existen diferencias entre, digamos, las percepciones de colores y las valoraciones. Hay muchas diferencias: sin embargo, por lo menos muchas de esas múltiples diferencias dejan intacta la propuesta específica de Hume, el punto específico de comparación al que estamos llegando ahora. Se trata de la idea de que, así como la atribución de cierta cualidad secundaria a los objetos ha de entenderse en el sentido de que esos objetos tienen una disposición en común para producir ciertas «percepciones en la mente», así también dentro del área conformada por los deseos potenciales del tipo

(4) la atribución de un valor específico a los objetos ha de comprenderse igualmente en términos de que esos objetos tienen una disposición en común para producir ciertas «percepciones en la mente» —una disposición en común, digamos, para producir valoraciones (de este tipo)—. Y la intención de la comparación sugerida en el contexto presente no consiste en sostener que tales valores son cualidades secundarias (así como la intención de rechazar la comparación sugerida no tiene que consistir en afirmar que tales valores son cualidades primarias): la intención es más bien la de liberar a la explicación dada aquí del área conformada por los deseos potenciales del tipo (4) de su apariencia de (cierto tipo de) «trascendentalismo». Por consiguiente, no va necesariamente en contra de la utilidad, en el contexto presente, de la comparación sugerida el hecho de que existan, por ejemplo, diferencias importantes entre los casos así comparados respecto de cómo han de entenderse las disposiciones implicadas —los tipos de diferencias que en parte llevaron a David Wiggins a hablar de «un subjetivismo de sujetos y propiedades *mutuamente* ajustados» (Wiggins 1987: 199) y de reacciones *«apropiadas»* de los sujetos (*ibid.*: 187), o también aquellas que en parte llevaron a John McDowell a hablar de «el carácter disputable que es típico de los valores», y de «objetos que *merecen*» la respuesta implicada (McDowell 1982: 119). (En esta discusión permitiré que el defensor de la propuesta de Hume juegue más desenfadadamente con la noción de disposición de lo que, según yo, tiene derecho a hacerlo (véanse las pp. 27–36), pues aquí mi tesis principal se manifestará con mayor claridad dentro del alcance de esta concesión.)

Aunque dudo que el sentido de «subjetividad» implicado positivamente en la propuesta a la que de tal modo hemos llegado se haya vinculado frecuentemente con la tesis de la subjetividad, por lo menos en boca de la gente normal, me parece sin embargo bastante claro que esa propuesta es compatible con la explicación dada hasta aquí de los casos de valoración dentro del área conformada por los deseos potenciales del tipo (4). Esa compatibilidad se garantiza mediante la combinación, dentro de dicha propuesta, del abandono de la tesis de la subjetividad en el segundo sentido de Hume y la introducción de una noción pertinente de disposición. No obstante, también me parece bastante claro que la propuesta a la que hemos arribado necesita un mayor desarrollo: en parte porque su carácter altamente abstracto significa que abarca demasiadas posibilidades distintas.

Volvamos, pues, una vez más a Hume. En su exposición, los valores (morales), tal como los colores, «no son cualidades en los objetos», sino

más bien los productos peligrosamente proyectables de «la constitución de [nuestra] naturaleza». Así que, para Hume, la comprensión filosófica de nuestro discurso sobre el valor (moral) debe enfocarse exclusivamente en esa «constitución» —es decir, debe enfocarse exclusivamente en las disposiciones pertinentes que se encuentran en nosotros, en las disposiciones pertinentes que hay *en los sujetos*—. Para el planteamiento al que acabamos de llegar aquí —humeano pero no de Hume—, el enfoque para comprender el sector pertinente de nuestro discurso sobre el valor en general parece inicialmente muy distinto, parece dirigirse en inicio a las disposiciones que hay *en los objetos*. Pero, desde luego, este enfoque no se dirige *exclusivamente* a las disposiciones que hay en los objetos: pues las disposiciones pertinentes se especifican en cuanto a los efectos potenciales que tienen «en la mente», en cuanto a sus efectos mentales potenciales *dentro de sujetos que tienen ciertas disposiciones*. Así pues, según el planteamiento al que hemos arribado, existen, *grosso modo*, dos tipos de disposiciones en juego: hay disposiciones en los objetos, las cuales se especifican en términos de sus efectos potenciales en ciertos sujetos portadores de disposiciones, y hay disposiciones en los sujetos, las cuales se especifican en términos de sus efectos potenciales en relación con ciertos objetos portadores de disposiciones. Y así se abre a la vista la posibilidad de que, debido a rasgos de los objetos y a rasgos de los sujetos, existan ejemplos particulares de los dos tipos de disposiciones que potencialmente se hallan en juego que no *encajan* bien unos con otros (véase Wiggins 1987a).

Así, de acuerdo con esta sugerencia humeana, la comprensión filosófica del discurso sobre el valor dentro del área conformada por los deseos potenciales del tipo (4) exigirá que se preste atención a dos tipos de disposiciones: disposiciones en los objetos, donde éstas se especifican, ciertamente, como disposiciones para producir, digamos, valoraciones de esos objetos en ciertos sujetos portadores de disposiciones, y disposiciones en los sujetos, donde se especifican como disposiciones por parte de los sujetos para valorar ciertos objetos portadores de disposiciones. Y ahora puede apreciarse el carácter altamente abstracto del planteamiento humeano con sólo enfrentarnos a una cuestión aparentemente simple: dentro de las especificaciones requeridas de las disposiciones en los objetos, ¿exactamente cómo ha de hacerse referencia a las disposiciones de los sujetos?

Al considerar esa cuestión, es importante recordar que el planteamiento de Hume se ha interpretado aquí como si se hubiera expuesto como una explicación descriptiva verídica de (nuestra comprensión de) nuestro dis-

curso sobre el valor dentro del área conformada por deseos potenciales del tipo (4). Pero consideremos ahora las siguientes proposiciones dentro de esa área:

(a) La búsqueda de la verdad histórica sobre el Renacimiento del siglo XII sería digna de ser emprendida, sería digna de valor, aun cuando se diera el caso de que los seres humanos y otros seres similares no sólo no la valoraran de hecho, sino que de hecho no tuviesen siquiera ninguna disposición a valorarla.

(b) La preservación de la vida humana inocente o de otras vidas similares debería ser valorada, merece ser valorada, tengan o no los seres humanos y similares siquiera la disposición de valorarla.

(c) La tortura no debería ser valorada, en verdad debería ser desvalorizada, al margen de las disposiciones humanas al respecto.

Desde luego, el mismo que formula tales proposiciones valora las cosas en cuestión (o en el caso de la tortura, pone de manifiesto su desvalorización de la cosa en cuestión); pero el alcance de sus atribuciones de valor a través de los «mundos posibles» parece claramente ir más allá de los mundos en que esa valoración ocurre, por su parte o por la de cualquier otra persona. En verdad, ese alcance parece «trascender» incluso la existencia de las disposiciones a valorar los objetos en cuestión. Parece haber, si se quiere, un elemento de *necesidad* en el contenido de esos pensamientos valorativos.

Si lo anterior suscita la objeción de que no tendría *caso* expresar esos pensamientos valorativos dentro del contexto de un público carente de las disposiciones en cuestión, la respuesta que ha de darse se compone de dos partes. En primer lugar, el hecho de que no tenga caso expresar un pensamiento semejante no equivaldría a la falsedad o la falta de sentido del pensamiento mismo. En segundo lugar, un objetivo razonable y posible para expresar un pensamiento valorativo de este tipo, dentro del contexto mencionado, sería el de producir las disposiciones pertinentes en el público en cuestión *respaldando ese pensamiento con la presentación de razones en apoyo de la afirmación de que la caracterización en él expuesta constituye realmente una caracterización de deseabilidad.* (Recuérdese que estamos interesados en el área conformada por los deseos potenciales del tipo (4).) Además, ello arroja luz sobre otros tipos de discursos valorativos de esta área que necesitan ser tomados en cuenta para propósitos descriptivos: el tipo de discurso que aparece en la afirmación de que la gente debería estar dispuesta a valorar las cosas particulares en cuestión sin importar su

disposición real a valorar, y la clase de discurso que aparece en la afirmación de que ciertas disposiciones a valorar son valiosas, son por lo menos dignas de ser valoradas, sean cuales sean las disposiciones reales a valorar de los seres humanos y de seres similares.

Dentro del área conformada por los deseos potenciales del tipo (4), hay entonces ciertas formas distintivas de hablar sobre el valor que el defensor de la tesis de Hume necesita ajustar si quiere que resulte aceptable como explicación descriptiva verídica de (nuestra comprensión de) todo discurso valorativo dentro de esa área. Quizás haya alguna afirmación que quepa en la muy abstracta idea inicial humeana, que constituya de hecho un desarrollo posible *de ella*, y que sirva también para ajustar las formas distintivas de hablar sobre el valor que aquí se discuten. Puede ser que, por ejemplo, para quien busca ese ajuste requerido, resulte asequible una explicación apropiada del *tipo* de referencia que se hace dentro de las especificaciones de las disposiciones pertinentes en los objetos, trátese de una referencia a valoraciones reales, a disposiciones reales a valorar, a la totalidad de las disposiciones reales a valorar o a las disposiciones a valorar realmente posibles. Lo que está claro es que incluso si el ajuste requerido se manifiesta como algo técnicamente posible, es difícil que tenga probabilidades de serlo en términos que sirvan para conservar la insistencia característica del subjetivista en las valoraciones y los deseos humanos reales, ni siquiera en las meras disposiciones humanas reales a valorar y desear, como la *fuente* de todo valor. Una vez que se distinguen los dos tipos de disposiciones implicadas en la propuesta humeana —disposiciones en los objetos y disposiciones en los sujetos—, y una vez que, consecuentemente, se aprecia el carácter distintivo de las disposiciones pertinentes en los objetos —la necesidad de entender esas disposiciones en términos de dignidad o mérito—, entonces ya se ha puesto en claro la posibilidad de cierto elemento de «trascendencia». El interesado en salvaguardar la insistencia característica del subjetivista debería haber protestado mucho antes.

En este capítulo he intentado desarrollar una teoría del valor y de la valoración utilizando la explicación del deseo y sus variedades que se presentó en los capítulos precedentes. El punto de partida fue el escepticismo cauteloso de Philippa Foot sobre «lo que queremos decir cuando hablamos de que *alguna cosa* tiene valor, o incluso de que la valoramos, en tanto opuesto a quererla o a estar dispuesto a tomarse la molestia de conseguirla». No hay una respuesta breve a ese escepticismo: todo depende

del tipo de deseo implicado. Aun así, formulados en los términos más sucintos posibles, los resultados de esta investigación son los siguientes:

(1) Los deseos «generadores de razones» no ofrecen base para el discurso sobre la valoración, de suerte que la identificación del querer con el valorar es incorrecta cuando el querer es de este tipo.
(2) Con respecto a la otra subclase de los deseos no motivados, aquellos cuyos objetos se desean al azar, la identificación del querer con una manera de valorar es sustancialmente correcta una vez que se imponen algunas otras condiciones a ese querer; pero la verdad de esa identificación no ofrece base para un discurso impersonal sobre los valores.
(3) Dentro de la primera subclase de los deseos motivados, tales como los deseos de los medios para los fines (deseados), cualquier identificación simple del querer con el valorar es falsa: pues existe la posibilidad del valorar sin el querer. Quizás sea más importante que se da una explicación cognoscitiva de las valoraciones dentro del área conformada por los deseos potenciales de este tipo; en conexión con esto, en tales casos se puede introducir un tipo restringido de discurso impersonal sobre el valor. Estas afirmaciones son, no obstante, compatibles con la tesis característica del subjetivista de que los deseos reales son la fuente de todo valor.
(4) Dentro de la subclase restante de los deseos motivados, en la que se centran las discusiones de Nagel y McDowell, también es cierto que cualquier identificación simple del querer con el valorar es falsa y que se da una explicación cognoscitiva de las valoraciones dentro del área conformada por los deseos potenciales de este tipo. Más aún, en tales casos se puede introducir un nivel todavía más impersonal de discurso valorativo. Pero estas afirmaciones en relación con esta subclase restante de deseos motivados son incompatibles con la verdad de la insistencia del subjetivista en los deseos reales como la fuente de todo valor. La apreciación verídica del carácter de estos deseos descarta, desde un punto de vista descriptivo, no sólo la identificación del valorar con el querer, sino además el proyecto subjetivista más vago de «reducir» todo valorar a desear.

La explicación cognoscitiva de las valoraciones dentro del área conformada por los deseos potenciales del tipo (4) se articuló en torno a un grupo de elementos clave (deseos, caracterizaciones de deseabilidad, características o rasgos de deseabilidad, razones) y de ciertas relaciones entre ellos

(en especial, las de «suscitar razonablemente» o «dar lugar razonablemente a»). Pero para una plena comprensión de esa explicación, todavía es preciso abundar más, con miras a una articulación detallada, tanto en esos componentes como, de un modo especial, en sus relaciones pertinentes. Parte de lo que se requiere se encontrará en la siguiente parte de este ensayo, cuando la atención se enfoque más de cerca a la cuestión de la moralidad.

Una vez que se reflexiona en la clasificación apropiada de los deseos morales, las valoraciones morales y los valores morales dentro de los términos de esta teoría del valor y la valoración, es probable que inicialmente parezcan obvias dos ideas. Una consiste en que por lo menos algunas valoraciones morales se comprenden en los términos del área conformada por los deseos potenciales del tipo (4), mientras que algunas otras se comprenden en los términos del área conformada por los deseos potenciales del otro tipo de deseos motivados, (3). La otra idea es que algunas de las valoraciones morales entendidas en los términos del área conformada por los deseos potenciales del tipo (4) se comprenden además en los términos más específicos de las respuestas cognoscitivas a cosas que merecen deseo, a las cuales se les debe deseo, mientras que otras valoraciones morales del mismo tipo general se comprenden además en los términos más débiles de las respuestas cognoscitivas a cosas que son meramente dignas de deseo, que son meramente objetos razonables de deseo. Pero por obvias que parezcan inicialmente esas ideas, surgen algunas dificultades.

Por lo que toca a la primera idea, la del carácter cognoscitivo entendido de ciertas valoraciones morales, hay una dificultad importante que se deriva de un argumento que Hume diseñó para mostrar que *en realidad* esa comprensión de las valoraciones morales incorporaría una ilusión. Ese argumento se dirige de hecho a mostrar lo que en realidad ocurre; es compatible, por tanto, con la idea *descriptiva* ya mencionada acerca de nuestra *comprensión* de ciertas valoraciones morales. Pero si el argumento de Hume es un buen argumento, sin duda deberíamos tratar de cambiar por lo menos nuestra comprensión de esas valoraciones. Además —y esto es más importante—, en la medida en que ese argumento se ha infiltrado en la conciencia ordinaria, se nos presentan dos posibilidades: una consiste en que la aceptación de ese argumento repose incómodamente al lado de la adhesión continuada a la idea descriptiva mencionada; la otra es que esa aceptación haya desplazado a dicha idea. En el primer caso, la adecuación descriptiva *plena* de la idea mencionada se ve amenazada (y la moralidad se encuentra verdaderamente en un lío); en el otro, peligra hasta su adecuación descriptiva *parcial*. No quiero adentrarme en la

cuestión histórico-cultural de si el argumento de Hume se ha infiltrado realmente en la conciencia ordinaria, aunque algo de lo que se expone más adelante (en el capítulo 6) servirá para marcar distinciones importantes entre los posibles resultados de semejante infiltración. Más bien, en el capítulo siguiente me limitaré a examinar, y de hecho me veré llevado a rechazar, ese argumento de Hume; sobre la marcha se dará parte de la formulación detallada que requiere la explicación cognoscitiva de las valoraciones morales pertinentes.

La otra idea inicialmente obvia que se mencionó incluía la afirmación de que ciertas valoraciones morales, entendidas en términos cognoscitivos y no subjetivos, se comprenden además desde el punto de vista más específico de las respuestas cognoscitivas a cosas que merecen deseo, a las cuales se les debe deseo. Esto parece implicar que cualquier diferencia en tal valoración es siempre un objeto de discusión y crítica potencialmente razonable. Pero muchos defensores del «relativismo moral» han aducido argumentos teóricos —algunos descriptivos, otros revisionistas— con la pretensión de mostrar, en efecto, la falsedad de esa aparente consecuencia. Por lo tanto, en el quinto capítulo examinaré algunos argumentos representativos de los «relativistas morales» con el fin de mostrar que no muestran tal cosa. En el curso de ese examen, saldrá a la luz la importancia de poner atención en caracterizaciones de deseabilidad apoyadas en razones específicas para la comprensión filosófica del pensamiento y la práctica morales.

Algunas otras cuestiones acerca de las conexiones que las teorías del deseo y el valor hasta aquí desarrolladas guardan con el pensamiento y la práctica morales se consideran en la segunda parte de este ensayo: la cuestión, por ejemplo, del lugar de los deseos no motivados de los tipos (1) y (2) en ese pensamiento y esa práctica. Y también se tocan algunos rasgos desconcertantes del uso real que hacemos de la moralidad. Pero el objetivo rector de la segunda parte de este ensayo es muy general: trabajar hacia una identificación de la institución de la moralidad por medio de la identificación de los rasgos centrales de la teoría interna del pensamiento y la práctica morales.

SEGUNDA PARTE

El filósofo supone que el valor de su filosofía yace en el todo, en la estructura; pero la posteridad lo encuentra en la piedra que usó para la construcción, y que después se usa muchas otras veces para construir... mejor. Así, encuentra el valor en el hecho de que la estructura puede destruirse y, *no obstante*, conserva su valor como material de construcción.

FRIEDRICH NIETZSCHE

Los conceptos morales no se mueven *dentro de* un mundo sólido fijado por la ciencia y la lógica. Construyen, para diferentes propósitos, un mundo diferente.

IRIS MURDOCH

CAPÍTULO 4

HECHO Y ACCIÓN EN LA TEORÍA MORAL DE HUME

> Una de las señales características, aunque no una señal necesaria, de un filósofo verdaderamente grande es cometer un error verdaderamente grande; es decir, dar una forma convincente y de influencia perdurable a una de esas falsas concepciones fundamentales a las que el intelecto humano se inclina cuando se ocupa de las categorías últimas del pensamiento.
>
> P. F. STRAWSON

> Creo que nunca conocí una mente más perversa y más obstinadamente inclinada al perjuicio público.
>
> WILLIAM WARBURTON, refiriéndose a Hume

4.1 EL ARGUMENTO MAESTRO

El examen que hace Hume del pensamiento y la práctica morales abarca una serie de afirmaciones distintas relacionadas con él. Quiero considerar aquí una preocupación central presente en ese examen, preocupación que conduce a Hume a presentar un argumento fascinante de gran maestría filosófica. La preocupación se relaciona con la afirmación de que «las reglas de moralidad [...] son [...] conclusiones de la razón» (Hume 1888: libro III, parte I, sección I).

El examen que hace Hume de esa afirmación tiene lugar dentro de una imagen en que la mente aparece compuesta por lo menos de dos facultades distintas, la de la razón y la de la pasión. La razón constituye la facultad cognoscitiva o intelectual, la facultad del entendimiento, cuya expresión paradigmática quizás sea la de las creencias. La pasión es la facultad activa o conativa, cuya expresión paradigmática puede considerarse aquí que es la de los deseos y voliciones. Los estados cognoscitivos y los conativos son tipos distintos de estados: hay una nítida distinción, por ejemplo, entre creencias, por un lado, y deseos, por el otro. Pero también, y esto es crucial, los estados cognoscitivos y los conativos se encuentran por lo menos en una dirección *aislados de cierta manera*: ningún estado

cognoscitivo de un agente, por sí solo, puede «dar lugar a» un estado conativo en él. En palabras de Hume, «la sola razón no puede nunca producir una acción o dar lugar a la volición» (1888: 414 [1977: 616]).

Signifique lo que signifique esa doctrina del aislamiento, hay dos afirmaciones, ambas verdaderas, que aparentemente Hume no está comprometido a negar con ella. La primera es que un estado conativo puede «tener lugar a partir de» un estado cognoscitivo en combinación con otro estado conativo. Por ejemplo, mi deseo general de beber agua, en combinación con mi creencia de que el vaso que tengo frente a mí contiene agua, puede «dar lugar» al deseo de beber el contenido de este vaso. Por lo tanto, no necesariamente hay contradicción por parte de Hume cuando dice, por ejemplo, que «cuando tenemos la *expectativa* de un dolor o placer proveniente de algún objeto, sentimos una *emoción consiguiente* de aversión o inclinación, y somos llevados a evitar o aceptar aquello que nos proporciona dicho desagrado o satisfacción» (1888: 414 [1977: 616], las cursivas son mías). Quien sienta aversión o tenga una inclinación general relativa al dolor o al placer, puede verse llevado a una aversión o inclinación frente a un objeto particular por la creencia de que ese objeto le causará dolor o placer. (Aunque más adelante –pp. 137-139, 144-145– señalaré una salvedad al respecto.)

La segunda afirmación que aparentemente no es necesario que Hume niegue es que algunas ideas –por ejemplo, las de emociones como el orgullo– pueden requerir, para que se puedan aplicar con verdad a una persona, que en esa persona se den tanto un estado cognoscitivo determinado como un estado conativo determinado. A lo que Hume está aparentemente comprometido es a la posibilidad de desenmarañar los componentes de una idea como ésta, a la posibilidad de analizar, por descomposición, las condiciones para su aplicación con verdad en los diferentes conjuntos de condiciones para la aplicación con verdad de cada una de sus ideas componentes, una de las cuales sería la de algún estado cognoscitivo y otra la de algún estado conativo. En realidad, ni siquiera es necesario que Hume niegue que la aplicación con verdad de algunas de esas ideas exige que se dé una relación entre un estado cognoscitivo por sí solo y un estado conativo; lo que tiene que negar es que el estado cognoscitivo por sí solo «dé lugar» al estado conativo.

Hume esboza también una distinción dentro de las operaciones de la facultad de la razón. En un tipo de empleo, la facultad de la razón se despliega en la búsqueda de verdades necesarias, verdades que dicen lo que tiene que ser el caso. Según Hume, las verdades de este tipo versan

exclusivamente sobre las relaciones entre las ideas: la «provincia propia» del uso correspondiente de la facultad de la razón es «el mundo de las ideas». Además, el conocimiento que puede proporcionarnos la razón respecto de esta «provincia» es un conocimiento *a priori*: dado que tenemos o poseemos las ideas pertinentes, nos hallamos por ello bien ubicados para alcanzar un conocimiento *a priori* de las relaciones entre esas ideas, un conocimiento *a priori* de las verdades necesarias. Pero las cosas son diferentes para el otro tipo de empleo de la facultad de la razón. Al desplegar la razón de esta segunda manera, lo que buscamos son verdades sobre la «provincia» de las «realidades»: verdades acerca de los objetos del mundo, sus propiedades y, especialmente, sus relaciones causales. Para Hume, cualquier conocimiento que obtengamos sobre esas verdades será *a posteriori*: sólo a través de la «experiencia», a través de la investigación empírica, podemos llegar a obtener este tipo de información. Más aún, Hume parece haber presupuesto tácitamente que el conocimiento *a posteriori* sobre la «provincia» de las «realidades» tendrá como objeto sólo verdades contingentes, verdades acerca de lo que es de hecho el caso, aunque las cosas podrían haber sido de otro modo.

Pese a la presencia de muchos aspectos polémicos en la manera en que Hume marca la distinción entre los dos usos de la facultad de la razón, su distinción es fundamental. Dicho de manera muy simple, al leer las obras de muchos filósofos anteriores a Hume —e incluso de muchos escritores posteriores—, no resulta claro a qué uso de la facultad de la razón se refieren mediante la mera palabra «razón». Pero hay una enorme diferencia entre un estado autocontradictorio y uno de ignorancia con respecto a algún hecho, así como hay una enorme diferencia, por ejemplo, entre la afirmación de que un amoralista se contradice en cierta forma y la afirmación de que ignora algunas verdades empíricas.

No obstante, los argumentos de Hume para mostrar que las «reglas de la moralidad [...] no son conclusiones de la razón» se dirigen igualmente contra cualquier intento de fundar el pensamiento y la práctica morales en cualquiera de los dos usos de la facultad de la razón. Pues esos argumentos están destinados a mostrar que ninguna conclusión de la razón, ningún estado de la facultad de la razón, podría desempeñar en nuestras vidas el papel que de hecho desempeñan nuestras ideas morales. Me centraré aquí en uno de esos argumentos, que corresponde a una preocupación más específica que forma parte de la preocupación general de Hume en torno a la refutación de cualquier tipo de «racionalismo» dentro de la filosofía moral. El argumento está concebido para mostrar que «tampoco consiste la moralidad en ninguna *cuestión de hecho* que pue-

da ser descubierta por el entendimiento» (Hume 1888: 468 [1977: 688]),[1] y se concentra en el segundo uso, el uso empírico, de la facultad de la razón; sin embargo, debería resultar obvio cómo se puede generalizar esta forma de argumento de modo que se oponga a cualquier «racionalismo» moral.

La mayor parte de los materiales exhibidos en este argumento humeano contra lo que denominaré «factualismo moral» los proporciona el mismo Hume en el curso de sus tentativas por establecer las siguientes afirmaciones: primero, «que la razón por sí sola no puede ser nunca motivo de una acción de la voluntad» (1888: libro II, parte III, sección III), y segundo, que las «reglas de la moralidad [...] no son conclusiones de la razón» (*ibid*: libro III, parte I, sección I). Y aunque formulo el argumento en términos que parcialmente no son humeanos, incluso este distanciamiento terminológico es quizás menor de lo que superficialmente parece. Dicho esto, podemos ahora considerar el argumento, numerando sus premisas conforme las vamos introduciendo.

En primer lugar, (1) una cuestión de hecho o de verdad es un objeto posible de conocimiento: su condición como tal cuestión está asegurada, y a la vez agotada, por la circunstancia de ser un objeto posible de conocimiento. Como lo expresa Hume:

> La razón consiste en el descubrimiento de la verdad o la falsedad. La verdad o la falsedad consiste a su vez en un acuerdo o desacuerdo con relaciones *reales* de ideas, o con la existencia y los hechos *reales*. Por consiguiente, todo lo que no sea susceptible de tal acuerdo o desacuerdo es incapaz de ser verdadero o falso, y en ningún caso puede ser objeto de nuestra razón. (Hume 1888: 458 [1977: 675])

> Si el pensamiento y el entendimiento fueran capaces de determinar por sí solos los límites de lo justo y lo injusto, el carácter de lo virtuoso y lo vicioso, esto último debería: o encontrarse en alguna relación de objetos, o ser una cuestión de hecho descubierta por nuestro razonamiento. (*Ibid.*: 463 [1977: 681])

> [T]ampoco consiste la moralidad en ninguna *cuestión de hecho* que pueda ser descubierta por el entendimiento [...] Pero ¿es que puede existir dificultad alguna en probar que la virtud y el vicio no son cuestiones de hecho cuya existencia podamos inferir mediante la razón? (*Ibid.*: 468 [1977: 688])

1. Para uniformar nuestra traducción con la de las obras de Hume que citamos algunas veces en este contexto traduciremos *matter of fact* como *cuestión de hecho*, y no simplemente como *hecho*, como lo veníamos haciendo. [N. de los tt.]

En segundo lugar, (2) el conocimiento de cualquier hecho o verdad particular sólo requiere, en lo que respecta a la persona que posee ese conocimiento, que se encuentre en un estado cognoscitivo determinado, en un estado específico de la facultad de la razón o el entendimiento. Como lo expresa Hume en otra obra posterior (*Enquiries Concerning Human Understanding and Concerning The Principles of Morals* 1902:[2] 172 [1945b: 30]): «Lo que es inteligible, lo que es evidente, lo que es probable y lo que es verdadero sólo obtiene el frío asentimiento del entendimiento y, al satisfacer una curiosidad especulativa, ponemos fin a nuestras investigaciones.» Pero (3) ningún estado cognoscitivo de una persona «da lugar» por sí solo a un estado conativo; las expresiones de la facultad de la razón o del entendimiento «no tienen influencia sobre las afecciones ni pueden mover las fuerzas activas de los hombres» (Hume 1902: 172 [1945b: 30]). De aquí que el conocimiento de cualquier hecho o verdad particular no requiera que se dé en la persona en cuestión un estado conativo, ni que «dé lugar» por sí solo a la existencia de ese estado.

Ahora bien, (4) para que una persona actúe o tenga siquiera una razón para actuar, se requiere *por lo menos* que se dé en ella un estado conativo, se requiere *por lo menos* que sus «poderes activos» se pongan en movimiento.

> La finalidad de todas las especulaciones morales es enseñarnos nuestro deber y, mediante representaciones adecuadas de la fealdad del vicio y de la belleza de la virtud, producir los hábitos correspondientes e inducirnos a evitar los unos y a abrazar los otros. Pero ¿cómo vamos a esperar esto de las inferencias y conclusiones del entendimiento que por sí solas no tienen influencia sobre las afecciones ni pueden mover las fuerzas activas de los hombres? Ellas descubren las verdades, pero si las verdades que descubren son indiferentes y no producen deseo ni aversión, no pueden tener influencia sobre la conducta y el comportamiento [...]
>
> Si extinguimos todos los cálidos sentimientos e inclinaciones favorables a la virtud, todo disgusto o aversión por el vicio, y hacemos a los hombres totalmente indiferentes a estas distinciones, la moral deja de ser un estudio práctico puesto que carece de toda tendencia a regular nuestras vidas y acciones. (Hume 1902: 172 [1945b: 30])

Por lo tanto, ningún conocimiento acerca de un hecho o verdad particular por sí solo da ninguna razón para actuar a una persona; es preciso que se dé un estado conativo adicional para que la persona tenga dicha razón.

2. Traducido en dos partes: *Investigación sobre el entendimiento humano* 1945a, e *Investigación sobre la moral* 1945b. [N. de los tt.]

Y, repito, ese estado conativo adicional no puede «surgir» del estado de conocimiento por sí solo.

No obstante, (5) la aceptación de un juicio moral puede proporcionar a la persona que lo acepta, y por lo menos a veces así lo hace, alguna razón para actuar; en realidad, la aceptación de un juicio moral en ocasiones da lugar a la acción misma.

> Si no fuera porque la moralidad tiene ya por naturaleza una influencia sobre las acciones y pasiones humanas, sería inútil que nos tomáramos tan grandes esfuerzos por inculcarla: nada sería más estéril que esa multitud de reglas y preceptos de que con tanta abundancia están pertrechados los moralistas. La filosofía se divide comúnmente en *especulativa* y *práctica*. Y como la moralidad se incluye siempre en el segundo apartado, se supone que influye sobre nuestras pasiones y acciones y que va más allá de los serenos y desapasionados juicios del entendimiento. Y esto se halla confirmado por la experiencia ordinaria, que nos muestra a los hombres frecuentemente guiados por su deber y disuadidos de cometer alguna acción por estimarla injusta, mientras se ven inducidos a realizar otras por creerlas obligatorias.
>
> Por lo tanto, [...] la moral influye en las acciones y afecciones [...] La moral suscita las pasiones y produce o impide las acciones. (Hume 1888: 457 [1977: 674-675])

Se sigue que la aceptación de un juicio moral por parte de una persona no es, respecto de esta persona, únicamente que se dé dentro de ella un estado cognoscitivo, y que tal aceptación no es únicamente cuestión de que la persona haya arribado al conocimiento de un hecho o una verdad: «Por lo tanto, las reglas de moralidad no son conclusiones de nuestra razón» (*ibid.*: 457 [1977: 675]).

Pero ahora, una cuestión de hecho moral o de verdad moral tendría que ser ni más ni menos que (i) una cuestión de hecho o de verdad, y, a la vez, (ii) eso solo, cuyo reconocimiento o conocimiento determina la aceptación por parte de una persona del juicio moral correspondiente de que el hecho o la verdad se da. De modo que una cuestión de hecho moral o de verdad moral tendría que ser tal que (i) el conocimiento de ella por sí solo no dé a una persona ninguna razón para actuar, y también tal que (ii) el conocimiento de ella por sí solo dé a una persona una razón para actuar. Conclusión: no hay hechos morales, la cuestión de la moralidad no es una cuestión de hecho o de verdad. Hume tiene la última palabra: «tampoco consiste la moralidad [...] en ninguna *cuestión de hecho* que pueda ser descubierta por el entendimiento» (1888: 468 [1977: 688]).

El argumento no pretende meramente refutar el factualismo moral, la idea de que hay hechos o verdades morales de algunos de los cuales, por lo menos, podemos saber que se dan, sino también elucidar el contenido de lo que niega. La idea clave es la conexión planteada en (1) entre las nociones de hecho y de conocimiento: para que una cuestión sea una cuestión de hecho no necesita ni más ni menos que ser un objeto posible de conocimiento. Esa idea se desarrolla algo más en la afirmación presentada en (2) acerca de, por así decirlo, la contribución constitutiva del cognoscente a su estado de conocimiento. Entonces, dadas las afirmaciones adicionales establecidas en (3) y (4) tanto sobre el aislamiento de un estado cognoscitivo por sí solo frente a los estados conativos de un cognoscente, como sobre lo que significa para una persona tener una razón para actuar, nos vemos llevados a la conclusión elucidatoria general de que ninguna cuestión de hecho o de verdad es tal que el conocimiento de ella por sí mismo da a una persona una razón para actuar.

La conexión entre las nociones de hecho y de conocimiento a partir de la cual se inicia esa elucidación general de lo factual no está libre de problemas; pero obsérvese que no es preciso que la conexión postulada ni la elucidación general final acarreen ningún compromiso con toscas teorías representativas de la adquisición de conceptos o del conocimiento. Consideremos de nuevo a Hume: «Lo que es inteligible, lo que es evidente, lo que es probable y lo que es verdadero sólo obtiene el frío asentimiento del entendimiento y, *al satisfacer una curiosidad especulativa, ponemos fin a nuestras investigaciones*» (1902: 172 [1945b: 30], las cursivas son mías). Algunas cuestiones de hecho pueden ser investigadas sólo debido a la «curiosidad especulativa» de los investigadores. En términos más generales, algunas cuestiones de hecho pueden ser sujetas a investigación sólo porque sucede que los investigadores tienen ciertas preocupaciones, intereses y necesidades: sin esas preocupaciones, intereses y necesidades, los investigadores no se habrían visto llevados a configurar los recursos conceptuales pertinentes ni a conseguir las evidencias pertinentes para el caso (véase Platts 1983a). Algunos estados no cognoscitivos de estos tipos (o de otros) pueden ser no solamente instrumentales para la adquisición del conocimiento en cuestión, sino esenciales para ella. Pero no es necesario que ello contravenga las doctrinas factualistas según las elucida el argumento de Hume. Si algún factualismo estuviera comprometido con el punto de vista de que los conceptos y la información pertinentes se nos imponen sin importar «ninguna otra consideración, cualquiera que ésta sea», difícilmente merecería ya que lo refutáramos.

Ahora puede verse que Hume adoptó una estrategia que comparten varios filósofos contemporáneos: la de abordar cuestiones metafísicas sobre hechos y valores mediante consideraciones de psicología filosófica. Dado lo infructuoso de la mayoría de los tratamientos «directos» tradicionales de esas cuestiones metafísicas, la atracción que ejerce este cambio de táctica resulta muy clara. Pero, en cualquier caso específico en que un filósofo haga uso de esta estrategia, es necesario considerar la cuestión de la adecuación de las secciones pertinentes de su psicología filosófica. Pronto volveremos a esto con respecto al caso de Hume; pero antes deseo terminar este breve esbozo de la parte de sus investigaciones sobre la naturaleza de la moralidad que aquí nos ocupa.

Dado el argumento que Hume ha presentado, ¿cómo puede explicarse el persistente apego a la tesis de que la moralidad es, en efecto, una cuestión de hecho o de verdad? Desde el punto de vista histórico, la más interesante de las explicaciones ofrecidas por Hume es la que se refiere a cómo la gente no reconoce la verdad de una premisa de su argumento: la premisa (4), según la cual para que una persona actúe o para que tenga siquiera una razón para actuar necesita *por lo menos* que dentro de ella se dé un estado conativo, necesita *por lo menos* que sus «poderes activos» se pongan en movimiento. La explicación de Hume sobre la incapacidad para reconocer esa verdad es la siguiente: hay ciertas «pasiones apacibles» que simplemente no se sienten, pues ni su existencia ni su eficacia causal pueden descubrirse mediante introspección (Hume 1888: 417 [1977: 621]); pues carecen del sentimiento que en otros casos es característico de las pasiones, se parecen, por lo tanto, exactamente a los estados cognoscitivos, a los estados de la facultad de la razón; pero luego, cuando esas «pasiones apacibles» actúan para producir efectos que toman la forma de acciones intencionales, parece, en consecuencia, que los estados de la facultad de la razón por sí solos pueden, en efecto, producir acciones intencionales. La razón por la cual digo que, históricamente hablando, ésa es la explicación más interesante que Hume da del persistente apego a la tesis pertinente es sencilla: vemos aquí el precio que Hume está dispuesto a pagar para sostener su teoría de la moralidad. Y el precio es elevado: admitir que existen «pasiones apacibles» lo obliga a abandonar uno de los principios fundamentales de su teoría de las ideas, el principio que sostiene que no podemos equivocarnos acerca del contenido de nuestra propia mente en un momento dado. El abandono consciente y sistemático de ese principio habría tenido consecuencias de gran alcance en otras partes de su filosofía. Y el hecho de que Hume no siguiera hasta el final

esas consecuencias no muestra que no se hubiera percatado de ellas, que no se hubiera percatado del precio que estaba pagando tácitamente para sostener su explicación de la naturaleza del pensamiento y la práctica morales.

Finalmente, Hume nos presenta, de una forma asombrosamente breve, sus propias ideas positivas sobre la naturaleza de la moralidad. A fin de explicar la naturaleza práctica del pensamiento moral, Hume cree haber mostrado que no se ganará nada con la postulación de propiedades adicionales de los objetos hacia los cuales se dirige nuestro pensamiento moral: cualesquiera de esas propiedades «objetivas» no serían más que otros hechos, otros objetos posibles de los estados de la facultad de la razón y, por consiguiente, dejarían enteramente inexplicada la presunta conexión entre el pensamiento moral y la acción. Más bien, lo que se necesita con el fin de explicar esto es algo *más* que hay *en nosotros*: algo que se relaciona adecuadamente con nuestra facultad pasional (véase Foot 1963: 74). Lo que Hume postula es un sentimiento distintivo de *aprobación moral*. Ese sentimiento es agradable o placentero; es más, Hume nos dice incluso que «no existe espectáculo tan hermoso como el de una acción noble y generosa, ni otro que nos cause mayor repugnancia que el de una acción cruel y desleal» (1888: 470 [1977: 691-692]). Hume cree que por este medio puede explicar nuestra tendencia a realizar acciones cuya contemplación produce un sentimiento tan delicioso; y por lo tanto cree que así puede explicar la naturaleza práctica de nuestro pensamiento moral.

4.2 La doctrina del aislamiento

Hume dijo que «la sola razón no puede nunca producir una acción o dar lugar a la volición» (1888: 414 [1977: 616]). Al lector contemporáneo quizás le resulte tentador interpretar esto como una mera expresión de Hume de ese aspecto de la ortodoxia corriente que sostiene que toda formulación plena de la razón para actuar de un agente debe hacer referencia no sólo a los estados cognoscitivos del agente, como es el caso de sus creencias, sino también a sus estados conativos, a sus deseos o «actitudes favorables» [*pro-attitudes*]. Pero si esto fuera todo lo que realmente significara la ortodoxia corriente, creo que habría que resistirse a la tentación. (En esta discusión, le obsequiaré a Hume la noción de razón para actuar, sin que esto quiera decir que él haya hecho libre uso de una noción semejante. En efecto, como pronto se pondrá de manifiesto, pienso que en la explicación oficial de Hume sobre la razón no hay lugar para nada

que se parezca a una razón para actuar. Pero el obsequio que le hago hace más fácil resaltar esa cuestión.)

Puede considerarse que Hume aceptaba por lo menos una parte de esta ortodoxia: aquella parte, sostenida en la premisa (4) de su argumento, que afirma que para que una persona tenga una razón para actuar se requiere *por lo menos* que se dé en ella un estado conativo. Su adhesión aparece quizás más nítidamente cuando invoca la idea de las «pasiones apacibles». Y no conozco ningún pasaje que indique que Hume hubiese negado el otro elemento de la ortodoxia corriente, según el cual tener una razón para actuar requiere que se dé un estado cognoscitivo como la creencia. Si la noción de tener una razón para actuar se vincula con la noción de la posibilidad de actuar conforme a la razón, tanto mejor, puesto que cualquier acción semejante necesita por lo menos una creencia por parte del agente sobre cómo, por ejemplo, buscar y alcanzar el objeto del deseo (véase Williams 1981a). Mi tesis consiste más bien en que ese aspecto de la ortodoxia corriente no agota el contenido de la afirmación de Hume.

Si eso era todo lo que Hume quería decir, entonces con la misma facilidad podría haber expresado su tesis diciendo que la *pasión*, por sí sola, no puede nunca producir ninguna acción ni dar lugar a la volición. Jamás lo expresó así. Y aunque sin duda había razones históricas específicas para que Hume se centrara en la impotencia motivacional de la razón, creo que el argumento humeano reconstruido revela que el asunto en cuestión no es sólo histórico. Lo que Hume necesita para los propósitos de ese argumento no es únicamente la parte mencionada de la ortodoxia corriente, sino además la afirmación de que ningún estado cognoscitivo de una persona «da lugar», por sí solo, a un estado conativo. Sugiero que esa adición forma parte de lo que Hume quería decir con su afirmación referente a la impotencia motivacional de la razón.

¿Qué implica ese añadido de Hume, la doctrina del aislamiento? ¿Y por qué habríamos de creerla? Hume escribe a menudo como si la tesis del aislamiento fuera meramente causal, como si la afirmación fuera que ningún estado cognoscitivo de una persona causa, por sí solo, un estado conativo. Si ésa es la tesis, entonces, entre todos los filósofos, Hume es el peor ubicado para sostenerla como una verdad *a priori*, porque una consecuencia importante de su discusión sobre la causación es que la determinación de los tipos de cosas que pueden estar en relaciones causales unos con otros no es nunca una cuestión *a priori*. Sin embargo, no hallamos nunca en los escritos de Hume los resultados de un examen cuidadoso de los

«objetos bajo una mirada filosófica estricta», concebido para apoyar su crucial añadido. Encontramos, en cambio, el siguiente argumento:

> Una pasión es una existencia original o, si se quiere, una modificación de existencia, y no contiene ninguna cualidad representativa que la haga copia de otra existencia o modificación. Cuando estoy encolerizado, yo poseo realmente la pasión, y en esa emoción no tengo una mayor referencia a algún otro objeto que cuando estoy sediento, enfermo o tengo más de 1.5 m de alto. Es imposible, por tanto, que a esta pasión se puedan oponer la verdad y la razón, o que sea contradictoria con ellas, pues esta contradicción consiste en el desacuerdo de ideas, consideradas como copias, con los objetos que representan. (Hume 1888: 415 [1977: 618])
>
> La razón consiste en el descubrimiento de la verdad o la falsedad. La verdad o la falsedad consiste a su vez en un acuerdo o desacuerdo con relaciones *reales* de ideas, o con la existencia y los hechos *reales*. Por consiguiente, todo lo que no sea susceptible de tal acuerdo o desacuerdo es incapaz de ser verdadero o falso, y en ningún caso puede ser objeto de nuestra razón. Ahora bien, es evidente que nuestras pasiones, voliciones y acciones son incapaces de tal acuerdo o desacuerdo, en cuanto que son hechos y realidades originales, completos en sí mismos, y no implican referencia alguna a otras pasiones, voliciones y acciones. Es imposible, por consiguiente, que puedan ser considerados verdaderos o falsos, contrarios o conformes a la razón. (*Ibid.*: 458 [1977: 675-676])

Uno de los pocos filósofos que han advertido la importancia de este argumento y que a la vez lo han interpretado explícitamente como un intento por defender la tesis meramente causal del aislamiento es Barry Stroud. Según esta interpretación, Hume está haciendo hincapié en la naturaleza abstracta de las entidades que son los únicos «objetos de la razón», es decir, las proposiciones. Siendo entidades abstractas, estos objetos de la razón no pueden tener influencia causal, pues esta influencia está restringida a las «existencias originales» o a las «modificaciones de la existencia». Pero luego Hume adopta la otra suposición «cuestionable» de que la razón en sí misma puede identificarse con la totalidad de los objetos de la razón. Y Stroud objeta que esto «parece dejar completamente de lado la noción de la razón como una facultad de la mente, o del razonamiento como un proceso mental» (Stroud 1977: 158-161 [1995: 227]).

La interpretación es ingeniosa; sin embargo, convierte este argumento de Hume en algo notablemente insignificante. El problema no es sólo el carácter cuestionable de la presunta suposición de Hume respecto de lo que la razón es. Aun concediéndole esto a Hume, su argumento, así interpretado, no hace nada por mostrar que un *estado* cognoscitivo de una

persona no puede causar por sí solo un *estado* conativo de esa persona. Pero, según la interpretación meramente causal de la tesis del aislamiento, la posibilidad aparente que «se deja completamente de lado» es exactamente la posibilidad que Hume tiene que negar para los propósitos de su argumento contra el factualismo moral.

Adviértase además que, según esta interpretación, dado que la «razón» no puede causar nada, ni siquiera es capaz de causar estados cognoscitivos como las creencias. Pero el argumento de Hume pone el énfasis en un supuesto rasgo distintivo de las pasiones: su falta de «toda cualidad representativa», la circunstancia de que son «hechos y realidades originales, completos en sí mismos». Pero, según la interpretación de Stroud, este énfasis resulta del todo innecesario: para esa interpretación, la impotencia causal de la «razón» es una verdad universal y la única que se necesita. Los comentarios sobre sus efectos (o la falta de ellos) no vendrían al caso.

(Algunos podrían pensar que hay una manera más rápida y directa de eliminar la interpretación meramente causal de la tesis de que ningún estado cognoscitivo «da lugar» por sí solo a un estado conativo, a saber, haciendo referencia a la discusión de Hume sobre el orgullo en el libro II del *Tratado*. Es cierto que Hume sostiene allí, en repetidas ocasiones, que la causa del orgullo es una creencia. Pero aparte de las conocidas dudas sobre la coherencia de los distintos libros del *Tratado*, pienso que Donald Davidson ha ofrecido la interpretación correcta de esas afirmaciones causales: «Ambas, creencia y actitud, razón y pasión, son necesarias para causar el orgullo. Pero las actitudes pertinentes [...] son, de acuerdo con Hume, universales [...] En lo que los hombres difieren es en sus dones y, por consiguiente, en lo que creen que son sus dones; esto es entonces lo que hay que mencionar cuando explicamos el orgullo, aun cuando se trate del principio "inactivo".» (Davidson 1976: 286 [1995: 356]))

Si ha de sacarse mayor provecho de Hume, resultará necesario, en mi opinión, abandonar la interpretación meramente causal de la tesis del aislamiento. Es decir, la cuestión de si un estado cognoscitivo puede causar por sí solo la existencia de un estado conativo tiene que distinguirse con claridad de la cuestión de si un estado cognoscitivo, por sí solo, puede «causar razonablemente» la existencia de ese estado conativo; en palabras de Hume, la cuestión de si un estado cognoscitivo puede dar lugar por sí solo a un estado conativo *en virtud de* alguna «conexión razonable» entre los contenidos proposicionales de esos estados.

Una posibilidad que surge para ser considerada gracias al establecimiento de esta distinción consiste en que Hume pensaba que no había

ninguna relación real del tipo de la causación razonable: la idea de tal relación es incoherente. El argumento para sostener esa idea podría ser, esquemáticamente, como sigue. Aunque tanto los estados cognoscitivos como los estados conativos poseen, en efecto, contenidos proposicionales, esos contenidos y sus relaciones de razón no tienen ninguna importancia para la cuestión de las relaciones causales que les conciernen, cualesquiera que sean. Si, por ejemplo, un estado cognoscitivo particular que es un estado de creencia causa alguna «modificación de la existencia» que es un estado de deseo, entonces, en lo que se refiere a la relación causal que se da entre esos estados, la cuestión de sus contenidos proposicionales es irrelevante —no tiene *nada que ver* con los poderes causales—. La eficacia causal es ciega a cualquier relación que la razón pueda detectar entre los contenidos proposicionales de esos estados: desde el punto de vista del nexo causal, el famoso cemento del universo, las «modificaciones de la existencia» no tienen más contenido proposicional que el hecho de que un hombre mida más de 1.5 m de estatura. Como lo expresa Hume:

> El entendimiento se ejerce de dos formas diferentes: en cuanto que juzga por demostración, o por probabilidad; esto es, en cuanto que considera las relaciones abstractas de nuestras ideas, o aquellas otras relaciones de objetos de que sólo la experiencia nos proporciona información. Creo que difícilmente podrá afirmarse que la primera especie de razonamiento es por sí sola causa de una acción. Dado que su ámbito apropiado es el mundo de las ideas, mientras que la voluntad nos sitúa siempre en el de la realidad, la demostración y la volición parecen por ello aislados uno del otro por completo. (Hume 1888: 413 [1977: 615])

De modo que el ámbito del razonamiento es el mundo de las ideas, mientras que la causalidad —así como la voluntad— nos sitúa en el de las realidades. De acuerdo con esta interpretación, Hume sostiene que la noción de causación razonable es la consecuencia incoherente del intento por unir esas dos provincias en una estructura federal.

Con esta interpretación, Hume está adoptando, en efecto, una de las respuestas eternamente populares al gran problema de cómo se puede conciliar el papel causal de los estados mentales con el papel racionalizador de los contenidos proposicionales de esos estados: Hume niega que tal conciliación sea posible y luego reacciona desechando la estructura ideal de la racionalización en favor del fundamento concreto y naturalista de la causalidad.

Quizás eso sea lo que Hume quería decir (dados otros elementos de su sistema filosófico); pero dudo que sea lo que realmente quiso decir. Según

esta interpretación, también debería haber negado que incluso un proceso a través del cual el razonamiento deductivo de una persona conduce a un cambio en sus creencias puede ser un proceso causal razonable, puede encarnar un cambio que se genera en virtud de una «conexión razonable» entre los contenidos proposicionales de las creencias en cuestión. Sospecho que Hume habría disfrutado esta consecuencia. Pero el asunto no acaba ahí. Según esta misma interpretación, Hume también debería sostener la misma negación respecto de la generación del deseo de salud a partir de la combinación de un aborrecimiento del dolor y la creencia de que la enfermedad es dolorosa. Esta generación tiene que ser entendida como algo meramente causal. Es más, tiene que sostenerse la misma negación acerca de la generación de intenciones y acciones por alguna combinación de creencias y deseos. Si mi deseo de beber agua y mi creencia de que esto es agua se combinan para generar mi acción intencional de beber esto, este proceso también tiene que ser entendido como meramente causal, como un proceso que no se produce de modo alguno a causa de los contenidos proposicionales de mis estados mentales.

Pero, en el mejor de los casos, eso resulta muy desconcertante. ¿Por qué ocurre que, de todas mis creencias, es de la creencia de que *esto es agua* de la que se asevera que se combina con el deseo para generar mi acción? ¿Es sólo porque de manera inverosímil pensamos que hemos observado una «conjunción constante» entre tales parejas de creencia-deseo, por un lado, y el beber agua, por el otro? Ciertamente no: para que la acción sea intencional según esa descripción tiene que generarse por medio de una pareja semejante, ya que sólo de ese modo se racionaliza de acuerdo con esa descripción. Ésta es una observación *a priori* sobre el concepto de acción intencional, sobre las condiciones en las cuales es posible que ese concepto tenga aplicación. Entendida correctamente, por lo tanto, no entra en conflicto con la afirmación de Hume acerca del desligamiento y la independencia de todos los seres en el universo (Hume 1888: 466 [1977: 685]). La concepción que la presente interpretación atribuye a Hume sólo puede sostenerse, por lo tanto, si se abandona el empleo del concepto mismo de acción intencional —concepto del que cabría esperar que figurara extensamente en cualquier tratado sobre la naturaleza *humana*—. (Adviértase además que alguien convencido de la interpretación que da Davidson de la discusión de Hume acerca del orgullo considerará que la presente interpretación se halla inmediatamente en conflicto con esa discusión en su conjunto. Y adviértase también que, según esta interpretación, Hume estaría comprometido a negar la primera de las afirmaciones citadas anteriormente (p. 126); sus observaciones acerca del papel de la

«perspectiva de dolor o placer» (Hume 1888: 414 [1977: 616]) tendrían que interpretarse desde un punto de vista meramente causal.)

Otra preocupación para quien defienda esta interpretación se encuentra en una cuestión textual relacionada que ya tocamos anteriormente. En ambos pasajes, Hume presenta su argumento como si se fundara en la naturaleza de las *pasiones* como «existencias originales», «modificaciones de la existencia», «hechos y realidades originales, completos en sí mismos». Sin embargo, según la interpretación que discutimos, Hume habría podido desarrollar su argumento con la misma eficacia asignando la misma naturaleza a los estados cognoscitivos, a los estados de la facultad de la razón (de ahí mis observaciones anteriores sobre el razonamiento deductivo). Es cierto que Stroud sostiene que el hecho de que una persona descubra o crea algo es «justo tan» «modificación de existencia» como el hecho de que esté enojada o de que mida más de 1.5 m de estatura. Pero seguramente es obvio que, aparte de toda cuestión meramente terminológica, Hume desea llamar la atención sobre algún rasgo que supone distintivo de las pasiones como fundamento de su tesis del aislamiento. La base de ese deseo es lo que todavía nos falta entender.

Obsérvese, por último, que no resulta nada claro cómo ha de modificarse, a partir de esta interpretación, el resto de la argumentación de Hume contra el factualismo moral, a fin de que sea válida y a la vez cuente con premisas verosímiles. La dificultad es obvia: para que la afirmación de que ningún estado cognoscitivo por sí solo causa razonablemente un estado conativo desempeñe la función que se le exige de llevar a la conclusión de que el conocimiento de una cuestión de hecho moral o de verdad tendría propiedades contradictorias, parece que resultará necesario invocar la noción de causación razonable dentro de una afirmación *positiva*, ya sea acerca de la noción general de tener una razón para actuar o acerca del carácter por el cual los juicios morales dan razones. Pero toda afirmación positiva semejante es *ex hypothesi* incoherente.

Volvamos a considerar las propias palabras de Hume cuando presenta el argumento en cuestión:

> Una pasión es una existencia original o, si se quiere, una modificación de existencia, *y no contiene ninguna cualidad representativa que la haga copia de otra existencia o modificación* [...] Es imposible, por lo tanto, que a esta pasión se puedan oponer la verdad y la razón, o que sea contradictoria con ellas, pues esta contradicción consiste en el desacuerdo de ideas, *consideradas como copias*, con los objetos que *representan*. (Hume 1888: 415 [1977: 618], las cursivas son mías.)

Digamos, de manera inicial y tentativa, que tanto el *creer que p* como el *desear que q*, por parte de una persona, son estados proposicionales de esa persona, cada uno de los cuales «se compone de» una actitud proposicional (creer y desear, respectivamente, en esos ejemplos) junto con cierto contenido proposicional (*p* y *q*, respectivamente). Así, quien cree que *p* no se encuentra por ello en el mismo estado proposicional que quien cree que *q*, ni quien cree que *p* se encuentra por ello en el mismo estado proposicional que quien desea que *p* (como se destaca con mayor claridad en los idiomas que exigen una nítida formulación subjuntiva de los contenidos de los deseos).

Ahora bien, para que un estado de creer que *p* cause un estado de desear que *q*, el estado de creer tiene que causar el estado de desear en virtud de alguna relación de razón entre los contenidos proposicionales *p* y *q*. Está claro, me parece, que Hume está tratando de caracterizar una diferencia fundamental entre, digamos, los estados de creer y los estados de desear, una diferencia que muestre la imposibilidad de un caso semejante de causación razonable. Pero, ¿cuál es exactamente esa supuesta diferencia?

Los estados de creer «contienen una cualidad representativa que los hace copias de otras existencias o modificaciones»; no así los estados de desear. Los primeros por lo menos pretenden representar cómo son las cosas, y por ello al menos se piensa que se «originan» en esa *presencia* que ellos *re-presentan*. En contraste, los estados de desear no pretenden representar cómo son las cosas, y por ello al menos se piensa que se «originan» en el interior de quien desea: son «hechos y realidades *originales* completos en sí mismos» (Hume 1888: 458 [1977: 675], las cursivas son mías). Ahora bien, Hume fue notoriamente negligente, por decir lo menos, acerca de la distinción entre conceptos y proposiciones y, en consecuencia, negligente, por decir lo menos, acerca de los modos en que puede determinarse el significado de una proposición mediante el significado de sus conceptos constitutivos. También parece haber pensado —dicho *muy* toscamente— que la cuestión de la significación de las proposiciones se determina «genéticamente», por las «existencias» que son causalmente responsables de que las proposiciones nos vengan a la mente. Esos dos rasgos del pensamiento de Hume pudieron haberse combinado con su concepción de que las pasiones son «existencias originales», de que se «originan» en nuestro interior, no afuera a partir de otras «existencias», para producir en él una inclinación a pensar que las pasiones carecen de contenido significativo, es decir, a pensar que no son realmente estados *proposicionales* en absoluto (véase Kenny 1963: 25, n. 1). Así, el mismo

Hume señala: «Cuando estoy encolerizado, yo poseo realmente la pasión, y en esa emoción no tengo una mayor referencia a algún otro objeto que cuando estoy sediento, enfermo o tengo más de 1.5 m de alto» (1888: 415 [1977: 618]). Pero entonces, si Hume pensaba que las pasiones como los deseos carecían de contenidos proposicionales, con toda corrección habría pensado que era obvio que ningún otro «ser en el universo» pudo razonablemente causarlas, de modo que habría pensado que era obvio que los estados cognoscitivos no pudieron hacerlo.

Esta interpretación encuentra mucho apoyo en los textos; yo creo que responde claramente a una inclinación que está en la mente de Hume. Pero, en primer lugar, le atribuye a Hume un error (hoy en día sorprendente): la negación a las pasiones de la condición de estado proposicional excluye toda posibilidad de explicar el papel racionalizador de las razones para actuar; por consiguiente, excluirá en última instancia el empleo del concepto de acción intencional. En segundo lugar, la interpretación se centra en un error innecesario que se descubre en cierto punto del sistema de pensamiento humeano, sin considerar de qué forma ese sistema global podía haberle permitido a Hume evitar el error. De ahí el comentario de Davidson sobre esta interpretación:

> La crítica es confusa. Lo que Hume *llamó* pasión carecía de «cualidad representativa»; pero sí la tenía el patrón de elementos que invocó para hacer de una pasión lo que es. La crítica válida consiste en que lo que Hume llamó pasión no tiene cabida en ese patrón. (Davidson 1976: 288)

En tercer lugar, resulta de nuevo oscuro cómo, de acuerdo con esta interpretación, ha de conservarse la validez del argumento de Hume. Y, finalmente, hay una interpretación de Hume más interesante desde el punto de vista filosófico, junto con un diagnóstico de su error, que tiene por lo menos el mismo grado de apoyo en los textos.

Quizás Hume no pretendía negar que las pasiones tuvieran contenidos proposicionales; más bien, lo que quizás pretendía negar es que tales contenidos fueran contenidos *representativos* o supuestas *copias*. Si se interpreta de esta manera y se dejan de lado complicaciones posteriores que aparecen en la discusión de Hume sobre la creencia, su tesis inicial se asemeja a la planteada por G.E.M. Anscombe en términos de las diferentes «direcciones de adecuación» que pueden darse entre los estados mentales y el mundo (Anscombe 1958: sección 2; y véanse, antes, las pp. 62–63). Un estado de creencia se propone representar cómo es el mundo, se propone adecuarse con el modo como el mundo es. Si falta tal «adecuación»,

entonces lo que debe modificarse para conseguir la «adecuación» necesaria es el estado de creencia. Los estados de creencia apuntan, *inter alia*, a la verdad, y es una falla decisiva en tal estado no acertar en el blanco. Pero los estados de deseo, como otros estados conativos, no se proponen representar cómo son las cosas, no se proponen «adecuarse» al modo como es el mundo. Si no se da esa «adecuación», entonces, para decirlo toscamente, quien desea no tiene ninguna obligación de modificar su estado de deseo a fin de lograr la «adecuación»; más bien, quien desea tiene razones para tratar de cambiar cómo es el mundo de modo que «se adecue» a (el contenido de) su estado de deseo. Mientras que un estado de deseo puede en apariencia dar en el blanco de la verdad, ésta no es nunca su meta. («En apariencia» debido a las complicaciones introducidas por el carácter subjuntivo de los contenidos de tales estados.)

Pero lo único que nos da eso es una elucidación de la distinción general entre estados cognoscitivos y estados conativos; hasta ahora, nada dice sobre las relaciones posibles, o imposibles, entre tales estados. Está claro que de dicha distinción no se sigue que ningún estado de creencia pueda causar, por sí solo, un estado de deseo, ni se sigue que los estados de deseo no sean en absoluto estados proposicionales. Con el propósito de salvaguardar una versión modificada del argumento humeano, lo que creo que se necesita es lo siguiente: la idea de que a partir de esta distinción entre estados de creencia y estados de deseo se sigue que ningún estado de creencia puede causar, razonablemente y por sí solo, un estado de deseo. Sin embargo, y esto es crucial, ello no será así por ninguna incoherencia en la noción relacional de causación razonable. Esa noción se considera coherente, y abre de este modo la posibilidad de una versión modificada del argumento de Hume contra el factualismo moral, en el cual se hace un uso positivo de dicha noción. Más bien, suponiendo que esa noción sea coherente, la idea es que se demuestra que carece de aplicación entre estados de creencia (por sí solos) y estados de deseo, en virtud de la distinción de tipo entre esos estados hacia la cual Hume llamó la atención. Pero, ¿por qué ese pensamiento podría parecer razonable?

En uno de sus pasajes más famosos, Hume escribió:

> No puedo dejar de añadir a estos razonamientos una observación que puede resultar de alguna importancia. En todo sistema moral que hasta ahora he conocido, he podido siempre observar que el autor sigue durante cierto tiempo el modo de hablar ordinario, estableciendo la existencia de Dios o realizando observaciones sobre los quehaceres humanos, y, de pronto, me encuentro con la sorpresa de que, en vez de las cópulas habituales de las

> proposiciones: *es* y *no es*, no veo ninguna proposición que no esté conectada con un *debería* o un *no debería*. Este cambio es imperceptible, pero resulta, sin embargo, de la mayor importancia. En efecto, en cuanto este *debería* o *no debería* expresa alguna nueva relación o afirmación, es necesario que ésta sea observada y explicada y que al mismo tiempo se dé razón *de algo que parece absolutamente inconcebible*, a saber, cómo es posible que esta nueva relación se deduzca de otras totalmente diferentes [...] una pequeña reflexión sobre esto subvertiría todos los sistemas corrientes de moralidad, *haciéndonos ver que la distinción entre vicio y virtud ni está basada meramente en relaciones de objetos, ni es percibida por la razón*. (Hume 1888: 469-470 [1977: 689-690], las cursivas en los dos últimos casos son mías.)

Nadie puede decir que los filósofos posteriores se hayan olvidado de este pasaje; pero me parece que no nos hemos preocupado lo suficiente acerca de su ubicación, contextual y filosófica, al final de los argumentos humeanos diseñados para mostrar que las distinciones morales no se derivan de la razón.

La interpretación de Hume sugerida comienza con la suposición de que, para él, la relación de causación razonable (de «dar lugar a») es un híbrido compuesto al menos por dos *tipos* de elementos completamente distintos: por un lado, relaciones de razón que se dan en el «mundo de las ideas» y que son candidatas para un conocimiento *a priori*, y, por el otro, relaciones causales que se dan en el mundo de las realidades y que son únicamente candidatas para un conocimiento *a posteriori*. Así pues, la causación razonable ocurre si se da una relación causal entre estados proposicionales en el mundo de las realidades, en virtud de alguna relación de razón que se dé entre los contenidos proposicionales de esos estados en el mundo de las ideas. Siendo esto así, bastaría en efecto una consideración de razón *a priori* para establecer en algunos casos que la relación de causación razonable *no* se aplica. (Cualquier verdad positiva que involucrara la relación sería únicamente un objeto posible de conocimiento empírico.) Es decir, desde este punto de vista de la relación, en efecto parecería posible determinar *a priori* que algunos tipos de cosas no podrían causar razonablemente otros tipos de cosas determinados.

Hume asegura conocer *a priori* que es «absolutamente inconcebible» que haya alguna relación razonable entre una proposición-*es* indicativa y una proposición-*debería*, de tal suerte que la segunda se siga de la primera. Esto es sólo cuestión de una «pequeña reflexión». Pero supongamos que Hume también consideró, como otros lo han hecho, que los deseos son consejos (dirigidos a uno mismo); o supongamos que pensó que una expresión natural de un deseo podía tomar la forma de una

proposición-*debería*; o supongamos que simplemente pensó que el contenido proposicional subjuntivo de un deseo se asemejaba mucho más a una proposición-*debería* que a una proposición-*es* indicativa. También pensaría entonces que su «pequeña reflexión» mostraría la imposibilidad de que un estado cognoscitivo cualquiera, con su contenido proposicional indicativo, causase razonablemente por sí solo un estado conativo. Dada la estructura de su argumento, pensaría por ello, en efecto, que su «pequeña reflexión» muestra «que la distinción entre vicio y virtud no es [...] percibida por la razón» (Hume 1888: 470 [1977: 690]).

4.3 Algunas dificultades

De acuerdo con esta interpretación, Hume se limita a suponer sin argumentación, en virtud de su uso de la noción de causación razonable, que las «relaciones ideales» de razón pueden reconciliarse de alguna manera con las realidades naturalistas de las relaciones causales. Sería injusto reprender a Hume por ese mero supuesto. Pero hay otras dos cuestiones por las que debe ser criticado.

En primer lugar, el asunto de su «pequeña reflexión» recibe un tratamiento demasiado simplista. No hay más que una mera afirmación acerca de lo que es «absolutamente inconcebible». Éste es un ejemplo de un defecto general de su filosofía, bien descrito y diagnosticado por Stroud:

> El tratamiento que da Hume al tema del «razonamiento a partir de meras ideas» es totalmente rudimentario y superficial. Puede aceptar, casi sin crítica, que se hable de «las mismas» y de «distintas» ideas, y de las «relaciones entre ideas», como si éstas habitaran en un dominio objetivo determinado inmediatamente abierto a la inspección de nuestras mentes. La teoría de las ideas alimenta esta representación de un conjunto de «cosas» autónomas, interconectadas, entre las cuales es posible descubrir que hay ciertas relaciones que subsisten meramente en virtud de que «poseemos» las ideas en cuestión. Constituyen una expresión de aquella representación las confiadas aseveraciones de Hume sobre cuáles ideas son «la misma» y cuáles son «distintas», con las consecuencias que tienen acerca de lo que es y lo que no es «absolutamente» imposible. (Stroud 1977: 240 [1995: 334])

Pero en el contexto presente resulta más importante otra crítica: la concepción de Hume del «mundo de las ideas» vuelve imposible el proyecto de reconciliar la razón y la causalidad. Según la representación de Hume, se juzga que la primera tiene como provincia propia el mundo de las ideas, la segunda el de las realidades; y es un mero juego de prestidigitación sugerir que los estados proposicionales pueden salvar de alguna

manera la brecha entre ambas provincias. Éstas ni siquiera podrían hallarse a tiro mental una de la otra. Para que se produzca un caso de causación razonable, como la entiende Hume, entre dos estados proposicionales, el uno causa el otro en virtud de alguna relación de razón en el mundo de las ideas entre las proposiciones que constituyen el contenido de esos estados. Pero, ¿qué significa aquí *en virtud de*? *No puede ser causal puesto que las relaciones de razón entre proposiciones en el mundo de las ideas, como las proposiciones mismas, son «seres» abstractos*. Pero, ¿qué podría haber sido para Hume si no causal? Concluyo que, para ser congruente, Hume tuvo que haber sostenido que la relación de causación razonable es incoherente (véanse las pp. 137–139); se aproximó a la razón por la cual tendría que haberlo sostenido cuando afirmó que las provincias de la razón y de la voluntad parecen ámbitos «aislados uno del otro por completo» (Hume 1888: 413 [1977: 615]). (Por tal razón tengo que retirar ahora el obsequio que inicialmente hice a Hume de la noción de una razón para actuar (véanse las pp. 133–134).)

Para Hume, las conexiones racionales y las conexiones causales se encuentran tan alejadas entre sí que resulta imposible darle sentido a la causación razonable como algo que acontece en el mundo natural. Más específicamente, resulta imposible darle sentido al *razonamiento* como algo que acontece en el mundo natural. Se requeriría por lo menos otro libro para considerar las posibles reacciones a este problema y para defender una teoría general del razonamiento y la razón que lo eluda sin caer por ello en las dificultades del psicologismo. Por el momento sólo quiero enfatizar que mi discusión se ha centrado en el papel que desempeña en la generación de ese problema la imagen que tiene Hume de la razón como fuente de un conocimiento *a priori* de «hechos recónditos sobre misteriosas entidades suprasensibles que poseen una vida propia» (Stroud 1977: 245 [1995: 342]) sólo porque creo que las deficiencias de esa imagen se aprecian con mucha facilidad. Pero también sospecho que cualquier explicación adecuada de este asunto necesitará además modificar la imagen humeana del ámbito de las conexiones causales como conexiones brutas e insensatas, por así decirlo, ya que sólo de esta manera se podría rescatar la relación de causación razonable de la condición híbrida que se le asigna en la teoría de Hume, y sospecho que ese rescate constituye un prerrequisito para el éxito de la tarea teórica pertinente.

4.4 Causación razonable de deseos

Quisiera desarrollar y ejemplificar lo anterior un poco más, a fin de conjurar diversos malentendidos que podrían surgir en relación con los aspectos exactos de la discusión de Hume sobre la moralidad que quiero poner en duda. Pero antes dos salvedades. (a) En apariencia, Hume estaba preocupado, ante todo, por (cierta impotencia de) los principios que rigen el razonamiento deductivo e inductivo, los principios de la razón teórica, y pensaba que éstos cubrían todas las tesis de la razón teórica; pero a mí me interesa más la posibilidad de que la «razón» tenga otro papel distinto. (b) Sin embargo, mi preocupación fundamental no será la conocida idea antihumeana de que hay principios de la razón práctica que Hume pasa por alto, principios requeridos por la racionalidad y que deberían gobernar la deliberación práctica. Es decir (combinando ambos descargos): si las nociones de racionalidad e irracionalidad se vinculan con las de obediencia o transgresión de los principios de la razón teórica o práctica, pronto se descubrirá que mi preocupación principal en este asunto no se refiere a cuestiones de racionalidad y de irracionalidad. Sin duda, hay casos de causación razonable potencial en los cuales el hecho de que no se produzca el «efecto razonable» potencial señala cierto elemento de irracionalidad por parte del sujeto implicado; pero en los casos de causación razonable que aquí me interesarán, no es preciso que ello sea así. En estos casos, la ausencia del efecto razonable potencial no basta para establecer la presencia de irracionalidad.

Una vez introducidas esas salvedades, pasemos a asuntos más positivos. En cualquier caso en que un conjunto de estados proposicionales razonablemente causa, «da lugar a», algún otro conjunto de estados proposicionales, el reconocimiento de esa causación razonable nos permite entender distintivamente por qué se produjo el efecto razonable. El cuestionamiento del *porqué*, interpretado de manera característica y natural, llega a un buen término (véanse las pp. 40–42, 71–73, 112–114). Sin embargo, repito, si en un caso estructuralmente similar el efecto razonable potencial no se produce de hecho, no es necesario concluir de ello que en este caso hay algún elemento de irracionalidad. Algunos de los casos relativamente poco problemáticos de esta causación razonable son los siguientes:

(1) Una pareja formada por una creencia y un deseo da lugar a una correspondiente intención de actuar.
(2) La intención de ejecutar un tipo de acción particular da lugar de hecho a una acción de ese tipo.

(3) La intención adicional con la cual podría ser ejecutado un tipo de acción da lugar de hecho a una acción de ese tipo.
(4) La creencia de un sujeto de que la acción de otro dirigida contra él manifiesta mala voluntad, da lugar al resentimiento en el sujeto.
(5) La creencia de un sujeto de que cierta situación en que se encuentra es peligrosa da lugar en él al temor.

En estos casos, podemos entender por qué el efecto se produjo razonablemente; pero si en otros casos, similares en cuanto a su estructura, no se produce ese efecto, no es necesario concluir que en ellos el agente transgrede algún principio de la razón teórica o práctica. Aceptamos la razonabilidad de la relación entre los estados proposicionales implicados; aceptamos tácitamente la razonabilidad de la relación distinta entre los contenidos proposicionales de esos estados proposicionales. Pero esa última razonabilidad no necesita reducirse a ninguna cuestión de *principios* de la razón teórica o práctica. Queda por resolver la cuestión de si en algún caso particular en que no ocurra el efecto razonable potencial existe algún elemento de irracionalidad por parte del sujeto; esto dependerá de la cuestión de cómo ha de explicarse la ausencia del efecto razonable potencial, pero no realizado. Para mis propósitos, basta con que las explicaciones en diversos casos semejantes puedan dejar intacto el problema del apego del sujeto a principios de la racionalidad, puedan dejar intacto el problema de la racionalidad de ese sujeto.

Pasemos ahora a otro tipo de casos. La persona *A* acaba de sufrir la muerte de su amada esposa; *B* y *C* son amigos íntimos de *A*, pero mientras que *B* tiene el deseo fuerte y manifiesto de ayudar a *A* en todo lo posible y durante todo el tiempo necesario, *C* no lo tiene.

Si consideramos el caso de *B* con un poco más de detalle, podríamos formularlo en los siguientes términos (véanse pp. 68–73):

(1) Primero está el *deseo* de *B*: su deseo, digamos, de ayudar a *A* en todo lo posible y durante todo el tiempo necesario.
(2) Luego tenemos cierta *caracterización de deseabilidad* que *B* tiene del objeto de su deseo: por ejemplo, la de ayudar a un amigo desconsolado tanto como sea posible.
(3) Finalmente, están las *razones* que *B* podría ofrecer, si tal fuera su manera de ser, en apoyo de esa caracterización de deseabilidad.

Incluso tomando en cuenta el asunto de la manera de ser de *B*, cuáles sean las razones exactas que aduzca en apoyo de su caracterización de deseabilidad dependerán del contexto de conversación específico en que

se encuentre. Pero ahora supongamos que *B* está hablando con *C* y que *B* llega razonablemente a pensar que la falta de un deseo como el suyo por parte de *C* no se debe a la torpeza o a la timidez de éste, ni a que *C* sea meramente un amigo sólo en las buenas, ni a alguna frialdad general en su carácter, ni a que tenga una concepción distintiva, más débil, de lo que la amistad es y lo que exige, sino más bien a que *C* tiene una opinión distintiva, y no poco común, de cómo vive la situación el acongojado *A*. Quizás *C* crea que la situación de *A* es perfectamente comparable, en cuanto a sus rasgos sobresalientes, a la de alguien cuyo ser querido se ha ido a vivir a una tierra lejana; o quizás *C* crea que la situación de *A* es perfectamente comparable, en cuanto a sus rasgos sobresalientes, a la de alguien que acaba de atravesar un doloroso proceso de divorcio.

Las características de la propia historia de *C* podrían intervenir naturalmente en el hecho de que *C* conciba la situación de *A* en esos términos: nuestras mentes muestran, en efecto, una enorme propensión a esparcirse sobre las mentes de los demás. Pero sea esto así o no, de todas formas no es preciso que *B* piense mal de *C*. No obstante, si ése fuera su modo de ser, *B* podría tratar de cambiar la opinión que tiene *C* de la situación de *A*: es decir, *B* podría intentar cambiar la concepción de *C* tanto en lo tocante a lo que es la muerte como en lo tocante a los rasgos sobresalientes de la situación de una persona que ha sufrido la muerte de un ser querido. Ante las analogías utilizadas por *C* según las hemos descrito aquí, *B* podría tratar, por ejemplo, de transmitir a *C* una estimación más vívida de lo que es vivir con una emoción intensísima dirigida a una completa ausencia *en* el mundo. No es muy probable que una tentativa semejante asuma la forma de un argumento derivado de los principios de la razón o la de un argumento fundado en los resultados de la investigación científica, y la cuestión de la probabilidad de que *B* tenga éxito en su intento tampoco es separable de cuestiones generales acerca del modo de ser de *C*; sin embargo, nada de ello demuestra que lo único de que se dispone en un caso semejante es la mera manipulación psicológica. (Con todo, algo que *se asemeja* a la investigación científica podría ayudar a *C* en ciertas situaciones específicas. En su investigación sobre los vínculos entre la tensión y la enfermedad (*Getting Well Again*), Simonton, Matthews y Creighton describen los esfuerzos de Holmes y Rahe para diseñar una escala que asigne valores numéricos a sucesos que ocasionan tensión. El fallecimiento del cónyuge se ubica en primer lugar con 100 puntos, el divorcio en segundo con 73: pero la «distancia» entre ambos se aproxima a la existente entre el último y la reconciliación marital —¡a la que se le asigna 45!—.)

La preocupación de *B* no es, desde luego, simplemente corregir una inadecuación cognoscitiva por parte de *C*, ni sólo corregir la falsa concepción de *C* respecto de la situación de *A*. Su preocupación consiste más bien en conseguirlo de tal modo que *en consecuencia* se produzca un cambio en los estados conativos de *C*, un cambio en los deseos de *C*. Si en realidad ocurre este cambio ulterior, podemos comprender por qué ocurre en cuanto al cambio del punto de vista de *C* sobre los hechos concernientes a la situación de *A*. Podremos comprender el deseo recién descubierto por *C* como un efecto razonable del cambio en sus estados cognoscitivos. Así es en efecto, creo, como lo entenderíamos de manera natural. No obstante, esto no tiene por qué implicar que si el cambio cognoscitivo ocurriera sin suscitar el cambio conativo, *C* manifestaría por ello algún elemento de irracionalidad, manifestaría por ello alguna transgresión propia de los principios de la razón. Comparemos los casos de emociones tales como el temor y el resentimiento: el hecho de que podamos entender la génesis de una emoción semejante en una persona en un momento dado desde el punto de vista de su concepción de los hechos pertinentes, ¿muestra que en casos estructural y cognoscitivamente similares en que no se genera esa emoción tiene que estar presente algún elemento de irracionalidad?

Es importante distinguir las dos observaciones hechas aquí. En primer lugar, he sostenido que a menudo comprendemos por qué diversos aspectos de nuestras vidas mentales se producen en términos que presuponen que hay relaciones de causación razonable que se encuentran entre ciertos grupos de estados cognoscitivos por sí solos y ciertos estados conativos. Y cuando se consideran las relaciones razonables entre los contenidos proposicionales de esos estados proposicionales, sostengo que creemos que hay relaciones razonables entre ciertos conjuntos de proposiciones-*es* indicativas y ciertas proposiciones subjuntivas que podrían acaso expresarse naturalmente mediante afirmaciones-*debería*. Esa primera tesis se vincula, pues, con las consecuencias de un punto de vista alternativo al que Hume sustenta sobre la razón y el razonamiento; pero esas consecuencias nos llevan a la segunda tesis, la cual apunta a un error relacionado que se encuentra en la teoría humeana de la motivación. Hume admite una manera en que la corrección de una creencia acerca de un hecho puede originar un cambio en los deseos de un agente, una manera en la que los estados de la facultad de la razón pueden afectar los estados motivacionales del agente. A muy grandes rasgos, la posibilidad que Hume admite puede surgir cuando un deseo más general se ha combinado con alguna falsa creencia factual para dar lugar a un deseo más específico.

Tenemos motivos para dudar de si Hume, para ser congruente, debería haber admitido esta posibilidad (pp. 144–145); pero la tesis presente es de distinta índole. Una vez que se entiende la primera tesis antihumeana aquí planteada, se reconocerá otra posibilidad —con el mismo carácter antihumeano— en el terreno de la motivación: los deseos de un agente pueden modificarse de manera razonable mediante la corrección de algún error en sus estados cognoscitivos, y ese cambio razonable entre los estados proposicionales constituye el «equivalente causal» de una relación razonable entre los contenidos de los estados cognoscitivos por sí solos y los contenidos de sus deseos —constituye el «equivalente causal» de una relación razonable entre cierto conjunto compuesto exclusivamente por proposiciones con *es* y ciertas proposiciones subjuntivas o que utilicen *debería*—. Intenté ejemplificar esta posibilidad con el caso en el cual la corrección de la falsa concepción de *C* sobre la muerte da lugar al cambio especificado de sus deseos; pero la existencia de este *tipo* de posibilidad es, desde luego, independiente (en cierta dirección) del carácter intuitivo del ejemplo dado.

Según la explicación del asunto que aquí se da, hay por lo menos tres maneras en que *B* podría no alcanzar su objetivo de modificar los deseos de *C* en relación con el desolado *A*: (i) *B* podría no lograr la modificación requerida de las opiniones que *C* tiene sobre lo que es la muerte; (ii) aunque *B* logre esa modificación, *C* no percibe empero la fuerza de la razón de *B* como una razón para aceptar la caracterización de deseabilidad pertinente; y (iii) si bien *C* llega a compartir la caracterización de deseabilidad de *B*, no llega empero a compartir el deseo de *B*. En el primer caso, el (i), *B* considerará que *C* permanece en un estado de ignoracia acerca de un hecho, y parece probable que *B* considere el segundo caso, el (ii), en los mismos términos. (Así será, sin duda, si el deseo potencial en juego pertenece, y *B* reconoce que pertenece, a la clase (4) dentro de la taxonomía del deseo presentada en el capítulo 2.) Y, por otra parte, tanto en el segundo caso como en el tercero, el (ii) y el (iii), *B* podría pensar que *C* ejemplifica el fenómeno de un efecto razonable potencial que no se produce en la realidad. Lo que hace que la noción de una caracterización de deseabilidad adquiera tanta importancia es el papel dual que desempeñan algunos de sus ejemplares en la estructura de ciertos casos: en estos casos, el fracaso en aceptar la caracterización de deseabilidad sería una manifestación de ignorancia acerca de un hecho, mientras que la aceptación de esa caracterización podría tener, como efecto razonable, el surgimiento de un estado conativo de deseo.

4.5 Moralidad y acción

Según el desaparecido J.L. Mackie,

> no hay indeterminación o falta de claridad en la tesis principal que [Hume] establece. Ésta consiste en que el hecho esencial, cuando la virtud se distingue del vicio o las buenas acciones de las malas, es sencillamente que las personas tienen sentimientos o sensaciones diferentes con respecto a ellos. (Mackie 1980: 64)

Pero así como lo obvio rara vez es obvio —lo obvio rara vez es *manifiesto*—, lo claro rara vez es claro. Creo que tras el «argumento principal» de Hume hay más de lo que se percibe a simple vista. (Desde luego, lo que aquí está en juego puede no ser más que una diferencia de intereses entre Mackie y yo. La afirmación de Mackie —que Hume supone que una actitud moral es una cuestión de sentimientos y que los juicios morales son expresiones de tales sentimientos— podría resultar bastante clara en el contexto de la oposición a los filósofos que no han empezado a apreciar la tesis.) Aquí no se considerará una gran parte del trasfondo oscuro; ni se considerarán las dificultades que surgen, por ejemplo, de la distinción de Hume entre virtudes naturales y artificiales, ni tampoco, de haberlas, de las opiniones de Hume respecto del carácter semántico de los juicios morales. Tampoco consideraré las ramificaciones de las siguientes observaciones hechas por Hume en sus escritos posteriores (aunque confieso que me habría encantado detenerme en su ejemplo de «una variación muy pequeña del objeto»):

> Si extinguimos todos los cálidos sentimientos e inclinaciones favorables a la virtud, todo disgusto o aversión por el vicio y hacemos a los hombres totalmente indiferentes a estas distinciones, la moral deja de ser un estudio práctico puesto que carece de toda tendencia para regular nuestras vidas y acciones. (Hume 1902: 172 [1945b: 30])

> Una pequeña variación del objeto, aun cuando se mantengan las mismas cualidades, destruye un sentimiento. Así, la misma belleza, transferida a un sexo diferente, no excita pasión amorosa cuando la naturaleza no está en extremo pervertida. (*Ibid.*: 213, n. 1 [1945b: 76])

> El celibato, el ayuno, la penitencia, la mortificación, la abnegación, la humildad, el silencio, la soledad y toda la serie de virtudes monásticas: ¿por qué razón son todas ellas rechazadas por los hombres de buen sentido sino porque no sirven para nada: ni aumentan la fortuna del hombre en el mundo, ni hacen de él un miembro más valioso de la sociedad, ni lo hacen capaz para el

> entretenimiento de las reuniones sociales, ni aumentan el poder del regocijo consigo mismo? (*Ibid.*: 270 [1945b: 137])

Quiero concentrarme, más bien, en «el aspecto principal» de la solución que Hume propone a *su* problema acerca de las conexiones entre moralidad y acción. En términos generales, esa solución es, repito, la siguiente: débase a la naturaleza o al artificio, el hecho es que contemplar ciertos tipos de acciones nos proporciona un tipo especial de placer, mientras que contemplar otros tipos de acciones nos causa un tipo especial de dolor. En consecuencia, tenemos una disposición natural a llevar a cabo acciones de los primeros tipos y a evitar las pertenecientes a los últimos tipos.

Quizás lo primero que debe destacarse es que Hume sostenía otra creencia que *requería* que su explicación de las presuntas conexiones entre moralidad y acción adoptara esa forma general. Hume aceptaba lo que Barry Stroud ha denominado «una doctrina monolítica de la motivación», según la cual cualquier cadena de deseos termina en el deseo de placer o de evitar el dolor (Stroud 1977: 170 [1995: 239]). Y si toda motivación se funda en ese deseo natural, también la motivación moral se funda en él.

Nadie necesita negar que muchos deseos humanos tendrían que comprenderse en términos de placer o de evitación del dolor (véanse las pp. 74–76). Muchos deseos apetitivos, por ejemplo, son desagradables e incluso dolorosos cuando no se los satisface, y su satisfacción, por el contrario, resulta agradable o incluso placentera, en parte debido a la eliminación del desagrado o el dolor que acompaña al estado de insatisfacción. Pero como tesis universal sobre las motivaciones humanas, la doctrina de Hume resulta indefendible —piénsese en los deseos cuyo contenido se refiere a momentos posteriores a la propia muerte, o incluso en los deseos cuyo contenido incluye la condición de que uno no sepa que se realizaron—. Por lo tanto, esa doctrina no nos ofrece ninguna razón para sostener que la explicación de Hume de la base motivacional de la moralidad *tiene que* ser correcta. Sin embargo, quien convenga en este rechazo de la doctrina monolítica de la motivación de Hume podría todavía afirmar que el resultado de aplicar esa doctrina al caso particular de la motivación moral ofrece en verdad la explicación correcta de ese caso como cuestión de mero hecho. Pero hasta esta afirmación se topa con dificultades, y entre éstas, una de las principales, y no teórica, es que choca con una intuición notablemente firme. Supongamos que, en efecto, siempre experimentamos un intenso placer al observar casos de cierto tipo de acción moralmente buena; no obstante, la intuición consiste en que la

perspectiva de experimentar ese placer no podría constituir la motivación *moral* para ejecutar una acción de ese tipo. La perspectiva de ese placer bien podría considerarse un grato *efecto secundario* de la ejecución de esa acción; pero si tal perspectiva constituye la motivación para llevar a cabo la acción, la motivación no es una motivación moral. En *Las relaciones peligrosas*, Valmont decide en cierto momento tratar de vivir «moralmente», y más adelante describe los resultados de su experimento de la siguiente manera: «Me asombró el placer que se deriva de hacer el bien, y me siento ahora tentado a pensar que las personas a quienes llamamos virtuosas tienen menos mérito del que se nos induce a creer» (citado en Dent 1984: 44, n. 4). La observación de Valmont es un arma de dos filos: registra el hecho de que una explicación al estilo de la de Hume podría servir para explicar muchas acciones *aparentemente* morales, y, al mismo tiempo, expresar la intuición de que, precisamente a causa de ello, las acciones así explicadas carecerían de mérito moral. (Esa intuición queda intacta ante las dudas que bien podrían suscitarse acerca de la veracidad de Valmont —quien, por jugar, decide entonces adoptar la moralidad— incluso como modelo de agente «moral» hipócrita.) O consideremos la anécdota de Aubrey sobre un sacerdote anglicano que, afuera de la catedral de San Pablo, vio a Hobbes dar limosna a un menesteroso y le preguntó si le habría dado la limosna si Cristo no lo hubiese ordenado así. Hobbes respondió que daba la limosna no sólo porque así complacía al pobre, sino porque lo complacía a él, a Hobbes, ver complacido al pobre. La congruencia teórica de Hobbes resulta evidente, pero la sospecha de que era un hombre moral trae consigo la sospecha de que estaba mintiendo. (El episodio se cita en MacIntyre 1966: 135.)

Esto no es más que la expresión de una intuición que se opone a la explicación que da Hume de la motivación moral; habría adquirido mayor interés si Hume hubiera compartido dicha intuición. Hay una cuestión interesante que surge al preguntarnos por qué Hume insistió en el carácter distintivo de los placeres y los dolores que constituían para él el fundamento de las motivaciones morales. Una posibilidad es que, mediante esa insistencia, Hume esperaba distinguir la moral de otras consideraciones motivacionales; otra es que trataba de allanar el camino para incorporar en su teoría la antigua doctrina de que las consideraciones morales para actuar tienen mayor peso que los demás tipos de consideración. Pero una tercera posibilidad es que, dentro del contexto de su doctrina monolítica de la motivación, Hume esperaba marcar así una distancia entre las motivaciones morales y motivaciones de otra índole. Esto es, su explicación de

la motivación moral no concibe la motivación como búsqueda del mero placer, sino más bien como búsqueda de un tipo distintivo de placer; y por consiguiente, podría decirse, el agente moral se distingue del mero buscador de placeres. De esta manera, la explicación humeana tal vez intentaría conservar cierta concepción de la *dignidad* o la *sublimidad* de la moralidad —para expresarlo en términos nada humeanos—. Pero incluso dejando a un lado la cuestión del apoyo que ofrecen los textos para atribuirle a Hume una estrategia semejante, sigue estando poco claro que esa estrategia se concilie con la intuición aparentemente contraria. Con el fin de mostrar la «naturalidad» de la motivación moral, resulta esencial, en los términos de la explicación de Hume, que la generación del placer distintivo no dependa de que quien lo experimenta piense que, por así decirlo, se ha realizado lo moralmente correcto. Sin embargo, en ausencia de esa adición vedada, sigue siendo cierto que la motivación moral se funda en la búsqueda del máximo placer natural, y a pesar de las posibles complejidades que acabamos de mencionar, esa afirmación sigue estando en conflicto con la intuición que aquí nos ha ocupado.

Subsiste el hecho de que una de las grandes contribuciones de Hume a la filosofía moral es el énfasis que puso en la naturaleza práctica de la moralidad, junto con su insistencia en la necesidad de examinar esa naturaleza conforme a los términos de una psicología filosófica aceptable. He tratado de indicar algunas de las razones que nos llevarían a pensar que la psicología filosófica en la que él mismo se apoya no resulta aceptable. Ahora tengo que reprenderlo respecto de otro tema.

Hablando en términos generales, Hume no se inclina a la repetición ni a la vaguedad cuando explica sus propias ideas. Pero sus descripciones y discusiones de la naturaleza práctica de la moralidad transgreden esta regla general. Los siguientes pasajes son buena muestra de ello:

> Si no fuera porque la moralidad tiene ya por naturaleza una influencia sobre las acciones y pasiones humanas, sería inútil que nos tomáramos tan grandes esfuerzos por inculcarla; nada sería más estéril que esa multitud de reglas y preceptos de que con tanta abundancia están pertrechados los moralistas [...] la experiencia ordinaria [...] nos muestra a los hombres frecuentemente guiados por su deber y disuadidos de cometer alguna acción por estimarla injusta, mientras se ven inducidos a realizar otras por creerlas obligatorias. (Hume 1888: 457 [1977: 674])

> La finalidad de todas las especulaciones morales es enseñarnos nuestro deber y, mediante representaciones adecuadas de la fealdad del vicio y de la belleza de la virtud, producir los hábitos correspondientes e inducirnos a evitar los

> unos y a abrazar los otros [...] Lo que es honroso, lo que es bello, lo que es decoroso, lo que es noble y lo que es generoso, toma posesión del corazón y nos incita a abrazarlo y a afirmarlo. (Hume 1902: 172 [1945b: 30])
>
> [L]a virtud heroica —siendo tan infrecuente—, es tan poco natural como la barbarie más brutal. (Hume 1888: 475 [1977: 697])

Así pues, el objetivo o propósito del pensamiento moral es modificar nuestro comportamiento mediante la modificación de nuestros «hábitos»; en ocasiones ese pensamiento logra su propósito, pero en otras no, mientras que rara vez se actúa conforme a las exigencias más extremas de la moralidad. Todo ello parece más o menos correcto; al mismo tiempo, nada de lo expuesto parece tener mucha utilidad para el esclarecimiento teórico de la naturaleza práctica de la moralidad.

No obstante, para ser justos con Hume, es preciso reconocer que en filosofía existe la dificultad general de producir una descripción de un fenómeno dado, con un nivel de exactitud interesante, en términos que no presupongan alguna teoría discutible del fenómeno en cuestión. Y esto hace que no resulte sorprendente que las descripciones de Hume de la naturaleza práctica de la moralidad, por lo demás anodinas, se complementen con las siguientes observaciones que, aunque tangenciales, son mucho más interesantes y mucho más teóricas:

> Tómese el caso de una acción reconocidamente viciosa: el asesinato intencionado, por ejemplo. Examínese desde todos los puntos de vista posibles, a ver si es posible encontrar esa cuestión de hecho o existencia a que se llama *vicio*. Desde cualquier punto que se mire, lo único que se encontrará serán ciertas pasiones, motivos, voliciones y pensamientos. No existe ninguna otra cuestión de hecho incluida en esta acción. Mientras se dediquen a considerar el objeto, el vicio se les escapará completamente. Nunca podrán descubrirlo hasta el momento en que dirijan la reflexión a su propio pecho y encuentren allí un sentimiento de desaprobación que en nosotros se levanta contra esa acción. He aquí una cuestión de hecho: pero es objeto del sentimiento, no de la razón. Está en ustedes mismos, no en el objeto. De esta forma, cuando reputan una acción o un carácter como viciosos, no quieren decir otra cosa sino que, dada la constitución de su naturaleza, experimentan un sentimiento o sensación de censura al contemplarlos. (Hume 1888: 468–469 [1977: 688–689])

Éste es uno de los pasajes que impulsaron a Mackie a señalar que «no hay indeterminación o falta de claridad en la tesis principal que [Hume] establece» (1980: 64); pero aún creo que tras la «tesis principal» hay más de lo que se percibe a simple vista.

Un replanteamiento tosco de las concepciones de Hume con respecto a la naturaleza práctica de la moralidad podría discurrir ahora como sigue. Cada vez que una persona acepta sinceramente un juicio moral sobre una acción ya ejecutada, tiene algún sentimiento, placentero o desagradable según el caso, que se dirige al objeto de ese juicio. En consecuencia, sucede que cada vez que una persona acepta sinceramente un juicio moral sobre una acción todavía no ejecutada, afronta la perspectiva de experimentar algún sentimiento, placentero o desagradable según el caso, perspectiva que se realizará si se da la circunstancia tanto de que la acción se ejecute de hecho como de que esa persona se percate de que se ha ejecutado. Frente a tal perspectiva, la persona tendrá por lo tanto un deseo correspondiente: el deseo de que se ejecute la acción, cuando el sentimiento previsto es placentero (junto con el deseo de percatarse de esa ejecución), el de que la acción no se ejecute, cuando el sentimiento previsto es desagradable (o el deseo de no percatarse de esa ejecución). Por lo tanto, lo cierto es que, dicho toscamente, cada vez que una persona acepta sinceramente un juicio moral acerca de una acción particular, tiene un deseo dirigido hacia la cuestión de la realización del objeto de ese juicio.

Deben hacerse algunos comentarios breves sobre este replanteamiento. En primer lugar, el contenido de estas afirmaciones de Hume sobre la naturaleza práctica de la moralidad depende de su concepción del deseo. Hemos visto que, al aceptar la existencia de «pasiones apacibles» (pp. 132–133), Hume abandona la idea, que por lo demás sería la idea empirista normal, de que los deseos se distinguen por la manera de sentirlos. También hemos hallado motivos para dudar de una interpretación ortodoxa de Hume, según la cual éste niega la condición de los deseos como actitudes proposicionales (pp. 140–141). ¿Qué nos queda, entonces, para distinguir los deseos de otras actitudes proposicionales? Parecería que lo único que queda —y hay abundantes pruebas textuales de su presencia— es la idea de que los deseos son disposiciones o propensiones a actuar, tendencias a tratar de obtener los objetos de los deseos. Pero en el primer capítulo hemos encontrado razones para dudar de la sustancia de toda idea semejante, y esa misma duda infecta, por lo tanto, la explicación más teórica de Hume sobre la naturaleza práctica de la moralidad. En segundo lugar, el replanteamiento expuesto rebasa todo lo dicho por Hume por lo menos de dos maneras: primero, el elemento del sentimiento se distingue en términos generales del elemento del deseo con el propósito de reducir la vaguedad del discurso sobre «sentimientos»; y segundo, el caso de una acción ya ejecutada se distingue del caso de una acción que todavía es meramente potencial con el propósito de revelar con mayor

claridad el aparato teórico que Hume despliega. (Desde luego, a pesar de su aceptación de las «pasiones apacibles», el pensamiento de Hume sobre el deseo se concentra, al igual que el de otros teóricos del poder activo (p. 83), en los deseos «generadores de razones»; en relación con esos deseos, la aplicación de la distinción general entre los elementos de sentimiento y de deseo resulta problemática.) Y el tercer comentario es que el replanteamiento difiere de todo lo dicho por Hume en su formulación del papel problemático que desempeña la conciencia que tiene el agente de la realización de la acción de que se trata. ¿Por qué no limitarse a desviar la mirada de las malas acciones de los demás? Sin embargo, vale la pena observar que descuidar ese papel resultará natural en alguien cuya única preocupación radique en los juicios morales que un agente tiene acerca de sus propias acciones.

4.6 Hecho y sentimiento

La creencia en una cuestión de hecho puede interactuar con algún deseo previo para dar lugar a un nuevo deseo. Mi creencia de que hay una botella de vodka en el refrigerador puede combinarse con mi deseo de beber vodka para producir el deseo de abrir el refrigerador. Pero esto no demuestra que la cuestión de que haya o no una botella de vodka en el refrigerador no sea una cuestión de hecho. En ocasiones, Hume se expresa como si ese modelo de la generación de nuevos deseos fuera el apropiado para la generación de deseos morales. Tenemos de modo natural un deseo general de placer y de evitación de dolor; ese deseo puede combinarse con la creencia de que, por ejemplo, sentiremos un placer característico al contemplar cierto tipo de acciones, para producir así el deseo de realizar acciones de ese tipo. Pero, aparte de la cuestión sobre el derecho que asiste a Hume para utilizar ese modelo (pp. 144–145), si ése fuera su ya considerado punto de vista del asunto, parecería que una consecuencia de ello es que la moralidad es, en efecto, una cuestión de hecho: una cuestión de qué acciones desembocan de hecho, al ser contempladas, en los placeres y dolores mencionados (véase Hume 1888: 468–469 [1977: 687], citado anteriormente, p. 155). La conexión requerida entre las creencias acerca de ese hecho y los poderes activos quedaría garantizada por el deseo general de placer y de evitación del dolor.

Para los propósitos de su argumento contra el factualismo moral, podría parecer que Hume necesita algo semejante a lo que sigue: la idea de que hay una conexión entre cualquier juicio moral sincero y un deseo correspondiente *independientemente de* otros deseos que el agente tenga.

Y tal vez la manera más humeana de tratar de fundamentar esa idea sea mediante la identificación, por lo menos parcial, del juicio moral y el deseo correspondiente. Aceptar con sinceridad un juicio moral *es*, por lo menos en parte, tener el deseo correspondiente independientemente de los otros deseos que uno tenga. Pero, ¿cómo puede conciliarse esta tesis con la doctrina monolítica de la motivación? El deseo general de placer y de evitación del dolor no puede desempeñar papel alguno en la *generación* de deseos morales si la moralidad no es una cuestión de hecho. Pero entonces, ¿cuál es la relación entre ese deseo general y la masa de deseos específicos implicados necesariamente en la aceptación sincera de los juicios morales? ¿Cómo podemos estar siquiera seguros de que esos deseos específicos no perjudicarán la doctrina monolítica de la motivación?

Un tanto irónicamente, podría parecer que la solución al enigma la han proporcionado los antihumeanos contemporáneos. Cuando Thomas Nagel introdujo la distinción entre deseos motivados y no motivados, su tercera explicación de los deseos motivados trazó una analogía entre su atribución y la atribución de creencias en principios generales de inferencia lógica (véase la p. 63). En esa ocasión, Nagel lo planteó de la siguiente manera:

> [L]a tentación de postular un deseo en la base de toda motivación se parece a la de postular que hay una creencia detrás de toda inferencia. Ahora podemos ver que la respuesta es la misma en ambos casos: que esto es cierto en el sentido trivial de que un deseo o una creencia está siempre presente cuando las razones motivan o convencen —pero no que el deseo o la creencia explique la motivación o la conclusión, o den una razón para ella—. Si alguien extrae conclusiones de acuerdo con un principio de lógica como el de *modus ponens*, es apropiado atribuirle la creencia de que el principio es verdadero; pero esa creencia se explica mediante la *misma* cosa que explica las inferencias que hace de acuerdo con el principio. (Nagel 1970: 30-31)

Aquí hay elementos en juego que Hume no puede aceptar sin caer en una contradicción; pero lo que, al parecer, sí podría aceptar es la distinción entre la atribución de un deseo general debido al papel que desempeña en la generación de otros deseos y la atribución de un deseo de esta clase debido a su «concordancia» o conformidad con un conjunto de deseos más particulares que se ponen de manifiesto en la conducta de una persona. Y si Hume sostuviera que la atribución del deseo general aducido en su doctrina monolítica de la motivación es del segundo tipo, podría

parecer que esta doctrina puede reconciliarse entonces con la negación de la tesis de que la moralidad es una cuestión de hecho.

Sin embargo, esta tentativa de reconciliación supondría un elevado costo. El problema no estribaría únicamente en la notable anticipación de la naturaleza implicada en el apego continuo de Hume a su doctrina monolítica de la motivación, ni en el hecho de que ese apego parece estar en pugna con los hechos de la total diversidad de los objetos del deseo humano, ni en el obstáculo que para esa doctrina representan casos reales de motivación moral, en conformidad con nuestra discusión anterior del caso de Valmont (p. 153). Ahora habría una dificultad distinta: si el intento de reconciliación aduce, del modo indicado, una noción análoga a la noción de Nagel de deseo motivado, es necesario mostrar cómo puede establecerse entonces la distinción entre el objeto del deseo y un previsible efecto secundario de su realización. Más específicamente: en los términos del intento de reconciliación contemplado, ¿cómo podría justificar Hume la idea de que el placer y la evitación del dolor constituye el objeto de ese deseo motivado, en contraste con el ser meramente el efecto secundario previsible de la realización de los objetos de distintos deseos particulares?

Sospecho que la verdad histórica relativa al propio punto de vista de Hume sobre el asunto es menos *recherché*. Hume escribió:

> Mientras se dediquen a considerar el objeto, el vicio se les escapará completamente. Nunca podrán descubrirlo hasta el momento en que dirijan la reflexión a su propio pecho y encuentren allí un sentimiento de desaprobación que en nosotros se levanta contra esa acción. *He aquí una cuestión de hecho: pero es objeto del sentimiento, no de la razón.* (Hume 1888: 468-469 [1977: 688-689], las cursivas son mías.)

> La moralidad es, pues, más propiamente sentida que juzgada [...] (*Ibid.*: 470 [1977: 691])

Tales observaciones provocaron un famoso reparo de Thomas Reid:

> Apliquemos este razonamiento al oficio del juez. En cualquier caso que se lleve ante él, el juez debe familiarizarse con todos los objetos y con todas sus relaciones. Después de esto, su entendimiento ya no tendrá más material en qué operar. No le falta, por su parte, nada más que sentir lo que está bien o lo que está mal; y la humanidad, muy absurdamente, lo ha llamado *juez*, cuando debería haberlo llamado *sentidor*. (Reid 1969: 474)

No me interesa aquí la fuerza filosófica de esa réplica tan agradable; el problema presente es de índole más histórica. Supongamos que las investigaciones de Hume sobre la naturaleza de la moralidad, al igual que muchas de sus investigaciones filosóficas, deban entenderse, por lo menos oficialmente, como si hubiesen sido emprendidas desde el punto de vista de la primera persona: esto es, como si hubiesen sido emprendidas por una persona que intenta comprender lo que sucede cuando, por ejemplo, ella misma hace juicios morales, *y cuya comprensión ha de evitar recurrir a las opiniones que otras personas tienen sobre el tema.* Supongamos además que, según tales términos, sus desaprobaciones y similares son objetos del sentimiento, no de la razón; esto es, tal vez sus sensaciones de desaprobación sean posibles objetos de conocimiento para otras personas y, en consecuencia, por así decirlo, constituyan *para ellas* hechos; pero desde su propio punto de vista, esas sensaciones son objetos del sentimiento, no de la razón ni del conocimiento y, por lo tanto, precisamente *para ella* no son hechos. Así pues, según los términos del punto de vista de la primera persona y del consecuente entendimiento de la noción humeana de «un hecho», en la explicación de Hume resulta que la moralidad no es en realidad «un hecho» —y resulta que no lo es en un sentido que es compatible con la doctrina monolítica de la motivación—. Por lo tanto, parafraseando un poco la afirmación de Mackie, podríamos decir ahora lo siguiente: *desde el punto de vista humeano,* lo esencial en la forma en que *yo* distingo la virtud del vicio, o las buenas de las malas acciones, es que *yo* tengo diferentes *sentimientos o afectos* al respecto. Pero el poco no es ni leve ni simple: ni la primacía ni siquiera la coherencia del presunto punto de vista así explicitado se encuentran a salvo de polémicas; la distinción exclusiva, dentro de este contexto, entre sentimiento y conocimiento es tendenciosa; y la noción operativa de «hecho» está ahora lejos de la noción, en apariencia más o menos común y corriente, que se nos presentó al principio de esta reconstrucción.

En la primera parte de este ensayo se desarrolló una teoría del valor que incluía un tratamiento cognoscitivo y no subjetivo de los miembros de cierta clase de valoraciones, y se sugirió entonces que, desde un punto de vista intuitivo, por lo menos algunas valoraciones morales pertenecían a esa clase. En este capítulo he examinado un argumento de Hume que, de ser válido, demostraría aparentemente la posibilidad de que dicho tratamiento cognoscitivo y no subjetivo de cualquier valoración moral sería correcto en el nivel descriptivo sólo al precio de revelar que esa valoración incorporaba una ilusión (véase la p. 121); pero he intentado mostrar

que el argumento no es válido. En particular, he criticado la premisa del argumento que afirma que ningún estado cognoscitivo puede «dar lugar» por sí solo a un estado conativo; en el proceso, espero haber articulado mejor los detalles tan importantes del tratamiento cognoscitivo y no subjetivo pertinente (pp. 146–150). Sin embargo, al examinar la explicación que da Hume de la naturaleza práctica de la moralidad presupuesta en su argumento, también he tratado de dejar claro el carácter muy especial del objetivo al que ese argumento se dirige; el resultado final, tal vez inesperado, es que aun cuando el argumento de Hume fuese válido, no perjudicaría *directamente* el tratamiento cognoscitivo y no subjetivo pertinente de las valoraciones morales.

Capítulo 5

EL ALCANCE DE LA MORALIDAD

> Los filósofos han medido montañas,
> penetrado en las honduras del mar, de los Estados y los reyes,
> han caminado con un báculo hasta el cielo, y rastreado las fuentes:
> pero hay dos cosas vastas y espaciosas
> cuya medición más importa, y sin embargo
> pocos hay que las sondeen: el Pecado y el Amor.
>
> George Herbert

> Bien dijo aquel gran hombre, creo que fue *sir* Walter Scott, pero si no él, fue el pequeño Bartley, el corredor de apuestas, que no existía un solo joven que no prefiriera una imputación a su moralidad antes que a su forma de cabalgar.
>
> R. S. Surtees

5.1 Las cuestiones internas

Thomas Nagel observó una vez que el problema del «*alcance*» de la moralidad constituye «uno de los más profundos en la teoría ética». Y añadió:

> Muchos han tomado como una objeción al utilitarismo el hecho de que obligue a la ética a engullirlo todo, dejando sólo una elección óptima o un pequeño conjunto de opciones igualmente óptimas como permisibles para una persona en un momento dado. Quienes ofrecen esta objeción difieren en cuanto a las dimensiones y la forma de la gama de opciones que deban quedar libradas a la inclinación individual después de haber fijado los límites éticos. (Nagel 1978: 116, n. 7)

Si en lugar de concentrarnos en el utilitarismo, nos concentramos en el tema, y por ende en la amenaza de que la ética lo engulla todo, peores habría, a guisa de primer paso, que el de distinguir tres cuestiones acerca del «alcance» de la moralidad en la vida de un individuo:

(1) ¿Cuál es el *área* de la moralidad? Es decir: ¿a cuánto de nuestras vidas son aplicables los conceptos morales? ¿Qué partes de nuestras vidas poseen, en potencia, una dimensión moral?

(2) Cualquiera que sea el área, así entendida, de la moralidad, ¿debemos tratar siempre de descubrir los aspectos morales pertinentes dentro de esa área? Dentro de la misma, ¿debemos siempre ir en busca de la dimensión moral potencial?
(3) ¿Cuál es el *dominio* de la moralidad? Después de haber registrado los aspectos morales de una situación específica, ¿debemos, por ejemplo, actuar siempre exclusivamente sobre la base de esas consideraciones morales? Dentro de su área de aplicación, como aquí se la entiende, ¿debe la moralidad ser soberana? ¿Cuándo deben prevalecer las consideraciones morales?

La pregunta sobre el área de la moralidad ha recibido con frecuencia una respuesta implícita a través de los ejemplos en que los filósofos han elegido concentrarse. Así, algunos filósofos se concentran exclusivamente en casos que asombran por mundanos y que a menudo resultan triviales; además, y esto es más importante, los casos están estrictamente circunscritos de acuerdo con los tipos de circunstancias en que pueden surgir. La recurrente obsesión, por ejemplo, por el pago de las deudas, el cumplimiento de las promesas, el acatamiento de las señales de tránsito. Nadie tiene que negar que se pueden encontrar aspectos morales en este tipo de ejemplos: pero si el área de la moralidad se restringe de esa manera —si ésos son los casos paradigmáticos para la aplicación del pensamiento moral—, resulta difícil entender entonces la preocupación que Nagel menciona: ¿cómo podría engullir mucho una moralidad semejante?

Filósofos de estilo muy diferente aparentemente consideran que las aplicaciones centrales del pensamiento moral se dan a la hora de tomar decisiones importantes y tortuosas. ¿Debemos bombardear Hiroshima? ¿Debo tratar de asesinar a Hitler? ¿Cómo debo comportarme en el *ghetto* de Varsovia? Nadie tiene que negar que se encuentran aspectos morales en este tipo de ejemplos: en realidad, resulta claro que en tales casos se suscitan profundas cuestiones morales. Pero si el área de la moralidad se restringe de esa manera —si ésos son los casos paradigmáticos para la aplicación del pensamiento moral—, entonces, por suerte, en la gran mayoría de las vidas humanas normales se corre escaso peligro de que la moralidad engulla algo.

La diferencia de estilo puede enmascarar la similitud entre las ideas involucradas en cada posición. Un aspecto de esa similitud está ahora claro: aunque de un modo distintivo, cada posición aísla a la moralidad por lo menos de buena parte de la vida humana. Pero debe observarse otro aspecto de la similitud: ambas posiciones se concentran en contextos

donde se toman decisiones para actuar. Nadie podría negar la idea de que la moralidad es práctica; pero está lejos de ser obvio que las relaciones del pensamiento moral con la práctica sean tan directas como esas posiciones las representan.

En *The Sovereignty of Good*, Iris Murdoch dio vívida expresión a una postura marcadamente diferente:

> El área de la moral, y por ende de la filosofía moral, puede entenderse ahora no como un asuntillo de deudas y promesas, sino como un área que cubre la totalidad de nuestro modo de vivir y la calidad de nuestras relaciones con el mundo. (Murdoch 1970: 97)

Hay sin duda muchas maneras de reconstruir los argumentos que condujeron a Murdoch a esa concepción del área de la moralidad (como aquí se la entiende); me limitaré a reconstruir el más sencillo. El pensamiento moral se dirige en realidad, por lo menos en parte, a cuestiones relativas a la acción, a cuestiones de la forma: «¿Qué haré?». Pero cuando se dan ciertas respuestas comunes a tales cuestiones —«Haz esto», «Eso es lo que deberías hacer», «Esto es lo que tienes que hacer», «Eso es lo mejor que puede hacerse»—, el interés y la sustancia de las respuestas depende de que estén potencialmente respaldadas por razones: razones que se expresarán o en términos de un vocabulario como el de los vicios y las virtudes tradicionales (lealtad, honestidad, valentía, etc.), o en términos de un vocabulario evaluativo aún más mundano (poco imaginativo, pedante, abusivo, etc.). Compárese con la observación de Wittgenstein de que sólo llamamos «hermoso» a un cuadro cuando no podemos molestarnos en pensar algo más específico que decir. Así mismo, en este sentido, los vocabularios relativos a vicios específicos y virtudes específicas y a las evaluaciones aún más mundanas, son los vocabularios primarios de la moralidad. Pero ahora debemos señalar, en primer lugar, que la aplicación de esos vocabularios no se limita a contextos en que se decide actuar, y en segundo lugar, que en una vida humana hay muy pocas cosas inmunes en principio a la aplicación de dicho vocabulario. La vida de una persona es casi totalmente una *posibilidad* perpetua de vicio y virtud, de valor y disvalor. Pero entonces, dada la primacía de los vocabularios correspondientes, el área de la moralidad es, en verdad, casi «la totalidad de nuestro modo de vivir y la calidad de nuestras relaciones con el mundo». No se puede escapar al alcance de la moralidad sólo dejando de pedir dinero prestado, negándose a hacer promesas y erradicando el uso del automóvil —ni siquiera en combinación con la suerte de no encontrarse nunca ante

un gran dilema—. (Aquí y a lo largo de todo este capítulo, se da al término «razón» un uso más amplio que el que hasta ahora ha recibido en este ensayo: ha de entenderse que abarca tanto lo que hasta ahora se ha llamado «caracterizaciones de deseabilidad» como lo que anteriormente se ha denominado «razones» (véase la p. 68). En el contexto actual esa distinción carece de importancia; más aún, según se pondrá de manifiesto, la mayoría de los ejemplos de lo que ahora se llamarán «razones» son ejemplos, en los términos anteriores, de «caracterizaciones de deseabilidad». Pero la naturalidad del uso que doy ahora al término «razones» supera, creo, cualquier ligero riesgo de confusión.)

Resulta peculiarmente difícil ofrecer un buen argumento en favor de una tesis que se impone como obvia. Sospecho que lo que en este caso ha ocultado esa obviedad es, con mucha frecuencia, una confusión entre la tesis citada sobre el área de la moralidad y una tesis similar sobre el dominio de la moralidad. La misma Murdoch nos dice, por ejemplo, que la filosofía moral «es el examen de la más importante de las actividades humanas» (1970: 78); más adelante observa que las artes «nos muestran la absoluta falta de sentido de la virtud a la vez que exhiben su suprema importancia» (*ibid.*: 86); así mismo, afirma que un «genuino sentido de la mortalidad nos permite considerar que la virtud es lo único valioso» (*ibid.*: 99); y concluye volviendo a señalar «la falta de sentido de la virtud y su particular valor y el alcance infinito de su exigencia» (*ibid.*: 104). Pero, de un modo tal vez más vívido para nuestros propósitos actuales, sostiene «que nada en la vida posee valor alguno, salvo el intento de ser virtuoso» (*ibid.*: 87). Buscando una expresión relativamente clásica y relativamente rigurosa de una concepción de fuerza similar, podríamos toparnos con Brentano: «Promover el bien a lo largo y ancho de este gran todo hasta donde sea posible; tal es evidentemente la meta correcta de la vida, y todas nuestras acciones deberían estar centradas en ella» (Brentano 1969: 32). Y vale la pena observar que hasta un escritor contemporáneo tan cauteloso como David Wiggins parece suponer que la cuestión de si la vida tiene un significado acaba por convertirse, por vías indudablemente sutiles, en la cuestión de si ciertas clases de juicios morales «pueden ser clara y aproblemáticamente *verdaderos*» (Wiggins 1976: 87).

Quizá en lo que sigue sólo exhibiré mi falta de comprensión de la noción de *virtud* de Murdoch, de la noción de *el bien* de Brentano y de la noción de *la moral* de Wiggins. De ser así, lo que en seguida diré habrá por lo menos de empezar a ponerlo en claro. Pero antes de eso, debería expresar algo cercano a la incredulidad ante estas afirmaciones:

incredulidad, por ejemplo, ante la idea de que la moralidad y la virtud desempeñen esos papeles excelsos en la idea misma de que la vida (o una vida) tiene significado, e incredulidad ante el aserto de Brentano acerca de lo que es *evidentemente* la meta correcta de la vida, en la que deberían centrarse *todas* nuestras acciones.

La propensión imperialista de una concepción de la moralidad aparentemente específica puede verse apoyada, desde luego de manera ilegítima, por la tendencia a expandir el alcance del concepto de moralidad: la tendencia a permitir que el *concepto* de moralidad engulla demasiado. Una vez entregados a esa tendencia, se perderá de vista la *singularidad* de la moralidad misma (para adaptar la útil expresión de Bernard Williams). Sin embargo, para luchar contra esa tendencia no será necesario ofrecer una caracterización o «definición» general de la «moralidad». Debe bastar, para los propósitos actuales, con que reconozcamos algunos ejemplos de diversos modos de pensamiento que, siendo prácticos, no son empero morales.

Unas semanas después de los terremotos de la ciudad de México de 1985, me hallaba viajando en taxi en la ciudad provincial de Guadalajara. Después de preguntarme acerca de la situación en la ciudad de México, el taxista me informó que los terremotos habían sido un castigo de Dios por el vicioso estilo de vida de los habitantes de la ciudad capital. No le pedí más detalles de su explicación, pero recordé la afirmación del emperador Justiniano de que la causa de los terremotos es la homosexualidad. No obstante, bajo las similitudes superficiales que había yo captado, había diferencias profundas entre las ideas de Justiniano y las de mi taxista.

Incluso para quien posee el concepto de moralidad, la pregunta «¿Qué debo hacer?» no tiene que ser equivalente a la pregunta «¿Qué debo hacer moralmente?»; en realidad, alguien que sepa la respuesta para la segunda pregunta puede, con toda coherencia, continuar preocupándose por la primera. Mientras que si alguien carece del concepto de moralidad, como Justiniano, cualquier supuesta equivalencia entre lo que se podría denominar «la cuestión práctica general», por un lado, y la cuestión moral, por el otro, no puede ser correcta. El juicio de Justiniano no fue y no pudo haber sido un juicio moral. Se pueden multiplicar los ejemplos: el preponderante respeto homérico por la astucia de un general, medida en función de su habilidad para engañar a sus oponentes; el código de honor que se manifiesta en la práctica del duelo; los dictados del machismo mexicano (o inglés). Ninguno de éstos es, ni pretende ser, un fenómeno que ejemplifique el pensamiento moral. Aunque esos ejemplos pudieran

parecernos cómodamente ajenos, no es esencial hacer referencia a ellos a fin de indicar la singularidad de la moralidad. Pues todos estamos ahora familiarizados con otros valores prácticos que pueden traerse a colación con respecto a la cuestión práctica general de «¿Qué debo hacer?». No es necesario poner gran énfasis en la cuestión de la prudencia —lo cual es afortunado, ya que el concepto mismo es un asunto de una oscuridad casi tan grande como el asunto del lugar que ocupan las consideraciones prudenciales dentro del pensamiento moral—. Los valores prácticos distintivos que se encuentran en la práctica de la etiqueta, en la ley, en relación con la conservación y la apreciación de las obras de arte, en las prácticas parcialmente constitutivas de la identidad cultural, e incluso, creo, en la política (recordemos la observación de Hume de que una buena filosofía política no es necesariamente una buena filosofía moral) ofrecen ejemplos mucho más claros. Junto a esa lista de valores relativamente «formales», podemos agregar otros, menos «formales»: aparte del asunto del humor, están por ejemplo los valores que se reflejan en «ideales individuales» particulares relativos a los estilos de vida (véase Strawson 1961), los que dan expresión y forma a caracteres y personalidades distintivos de los individuos, y los que surgen dentro de relaciones personales distintivas, como el amor y la amistad.

Cada una de esas cosas valoradas puede pesar sobre la cuestión práctica general de «¿Qué debo hacer?», y cada una de ellas puede continuar pesando incluso cuando la moralidad misma interviene en la cuestión. La tesis de que la moralidad por sí sola debería determinar siempre nuestra respuesta a la cuestión no es ni trivial ni obvia. Aun si pasamos por alto las patentes dificultades que se suscitan para el término *debería* tal como aparece en la tesis, tenemos que advertir, y rechazar, otra maniobra concebida para revelar la obviedad de la tesis, otra maniobra concebida para hacer evidente que la moralidad debería engullirlo todo: la maniobra que toma la forma de afirmar que los valores de todas esas otras cosas se derivan de los valores morales. Pensemos, por ejemplo, en la aserción de que la única razón que tenemos para conservar las grandes obras de arte es que tenemos una obligación moral de conservarlas. Cualquier aserción de este tipo sirve simplemente para asentar la determinación de tratar a la moralidad misma como un tragalotodo. El rechazo de esa determinación no *requiere* la negación de la idea de que la moralidad debería acaso mostrar algún tipo de respeto hacia los lugares que ocupan esos otros valores; podría simplemente hacer asentar que éstos poseen sus propios lugares.

Un último ejemplo de tentativa por esquivar las dificultades que hay aquí. En *A Theory of Reasons For Action*, David Richards propuso el siguiente principio:

> [U]n principio de amor mutuo que exija a las personas que no muestren afecto personal ni amor hacia otros únicamente sobre la base de unas características físicas arbitrarias, sino, más bien, sobre la base de los rasgos de la personalidad y el carácter que se relacionan con la acción fundada en principios morales [...] (Richards 1971: 94)

¿Cómo un hombre a todas luces inteligente pudo llegar a escribir algo semejante? Me parece que una parte de la respuesta es bastante clara: Richards advirtió que si desde el principio no incluimos dentro de la moralidad las relaciones personales como el amor y la amistad, luego esas relaciones pueden adquirir una fuerza o incluso autoridad, desde un punto de vista motivacional, capaz de amenazar la autoridad de la moralidad. De suerte que, en un intento por mantener esta autoridad de la moralidad —su dominio—, Richards trató de reinterpretar la naturaleza de esas relaciones personales de manera que esa amenaza no pudiera surgir jamás. Las relaciones personales han de verse como un tipo de relación moral; así, lo que la moralidad dice no entrará en conflicto con lo que dicen las relaciones personales. Pero antes de apresurarnos a abrazar este principio del amor moralmente purificado, primero tenemos que preguntarnos si en realidad queremos que la vida humana sea gobernada por un principio semejante. Y si se guarda alguna duda en ese respecto, sólo necesitamos recordar una «solución» alternativa: la de aceptar la pluralidad irreductible de los *tipos* de valores prácticos dentro de la vida humana, y la consecuente necesidad, en esa vida, de hacer frente a los conflictos entre dichos valores. (Ciertamente, en el planteamiento de Richards también está presente una concepción de las reglas como el concepto primario de la vida moral que, según Alasdair MacIntyre, es la postura característica de la modernidad: «Las cualidades del carácter [...] por lo general sólo llegan a ser apreciadas porque nos conducirán a seguir el conjunto correcto de reglas» (MacIntyre 1981: 112). Con respecto a esta cuestión, que en otros contextos es más importante, véanse también McDowell 1979 y Warnock 1971. El principio de Richards del amor mutuo me parece una *reductio ad absurdum* del casamiento de esa postura con la exageración del dominio de la moralidad.)

Los aspectos principales en que aquí hemos insistido han sido, en primer lugar, el carácter distintivo de la cuestión relativa al área de la moralidad

con respecto a la referente a su dominio y, en segundo lugar, la dificultad de esta última cuestión a la luz de la pluralidad de los valores prácticos. Quizá una vida sin moralidad sería solitaria, pobre, desagradable, brutal y breve; pero aún no está claro que una vida en la que la moralidad constituya una soberana todopoderosa dentro de su vasta área sería notablemente mejor. Habría que dirigir el pensamiento y la imaginación, por ejemplo, hacia el carácter de un modo de vida en el que las consideraciones morales prevalecen siempre sobre los ideales individuales de los agentes, sobre la expresión y el desarrollo de sus caracteres y personalidades, sobre las motivaciones emanadas de sus relaciones personales, como el amor y la amistad, el odio y la enemistad. No es sólo que no seamos lo que serían los practicantes de un modo de vida semejante; más bien sabemos que, si llegáramos a ser así, habríamos perdido demasiado. Demasiado de nuestro ser, del ser de los otros y de las relaciones entre estos seres. Una vez perdido todo eso, bien podríamos haber perdido todo: pues habríamos perdido realmente *nuestras* vidas.

Habría que mencionar otra posibilidad antes de dejar estas dificultades sobre el alcance interno de la moralidad. Podría plantearse la propuesta de que necesitamos distinguir cierta subclase privilegiada de los casos en los que la moralidad se expresa sobre la cuestión práctica general: «¿Qué debo hacer?» —es decir, los casos en que la moralidad *ordena*—. El contraste entre una *exigencia o demanda* moral y una mera *consideración* moral, contraste en que se basa lo anterior, parece que viene de todos modos exigido por nuestras ideas acerca de la supererogación y el heroísmo (véase Urmson 1958). Y ahora, prosigue aquella propuesta, deberíamos limitarnos a sostener que, en casos pertenecientes a esa subclase privilegiada, la voz de la moralidad es preponderante. La idea es que mediante esta estrategia podemos responder a la cuestión sobre el dominio de la moralidad; y no obstante, prosigue la idea o la esperanza, ese dominio no será, en general, más que una pequeña parte del área de la moralidad. Y así, concluye la idea, ya no habrá peligro de que la moralidad termine por engullir «la totalidad de nuestro modo de vivir» a expensas de todas las demás consideraciones evaluativas.

Llevar a la práctica esa estrategia constituye una tarea difícil. No hay motivos para creer que tenga que adoptar la forma de una caracterización general de la subclase privilegiada de casos, una caracterización general de qué es lo que hace que haya una exigencia o demanda moral, en la que la caracterización no emplea ella misma la idea de exigencia o demanda moral. (Esa forma se encuentra ejemplificada en el intento, independien-

temente poco atractivo, de poner en práctica la estrategia mediante un simple listado de las exigencias morales y de las circunstancias que las suscitan, como si se tratara de una enciclopedia de deberes codificados: pues cualquier intento así se presta a proporcionar una caracterización, por desaliñada que sea, de aquel tipo general.)

Pero entonces, ¿cómo precisamente se debe poner en práctica la estrategia? Una posibilidad sugerente se apoya en nuestra discusión anterior sobre la falsedad del punto de vista atomista de la percepción de la deseabilidad (pp. 55–57). Allí se afirmaba que a uno le puede parecer deseable cierto rasgo de una acción en tanto que solamente preste atención a ese rasgo, pero que podemos terminar por considerarlo totalmente carente de deseabilidad al colocarlo dentro del contexto de la acción como un todo. Eso podría interpretarse nada más que como la descripción de una realidad psicológica, de un hecho psicológico. Pero podría surgir ahora la siguiente idea: cuando existe una exigencia, algo que en otro caso habría sido una razón para actuar de manera diferente, que en otro caso habría constituido una consideración de deseabilidad, no es una razón en absoluto para actuar de manera diferente, no es una consideración de deseabilidad; de modo que, cuando una persona que contempla correctamente las cosas llega a advertir que algunos rasgos en apariencia deseables de otros cursos de acción posibles para ella lo son sólo en apariencia, está en juego una exigencia; por consiguiente, el agente que contempla correctamente las cosas dentro de un contexto semejante ejemplificará la realidad psicológica que acabamos de describir. Ésta es una observación acerca de un *efecto*, que se ajusta a la descripción general que acabamos de ofrecer de la realidad psicológica, dentro de la psicología de una persona que contempla correctamente las cosas. Pero esta última noción exige que la circunstancia de que una cosa sea o no una razón —de que una cosa sea realmente un rasgo deseable— sea, por lo menos en los casos pertinentes, un asunto objetivo, no constituido por hecho psicológico alguno. (Compárese con la discusión sobre la «subjetividad» de los valores en las pp. 114–116; pero véanse también las observaciones finales en las pp. 113–114.) Así, esta manera de vincular la noción de exigencia con la psicología de los agentes no requiere la idea de que cualquier efecto psicológico que se ajuste a la descripción general aquí ofrecida de la realidad psicológica pertinente pone de manifiesto la presencia de alguna exigencia: el error siempre es posible. En realidad, esta forma de vincular la noción de exigencia con la psicología de los agentes ni siquiera requiere la idea de que la aparición de la realidad psicológica pertinente pone de manifiesto la presencia *supuesta* de cierta exigencia: la descripción dada

de la realidad psicológica pertinente es, en efecto, general. Y ahora, por último, para completar la idea que estamos considerando, puede agregarse —sólo como un fragmento de fenomenología— que en los casos en los cuales la exigencia que está en juego es moral, el agente que contempla correctamente las cosas y, en consecuencia, experimenta el efecto psicológico descrito, experimenta por ello la «sublimidad» de las exigencias morales.

Muchos lectores habrán querido protestar desde el principio por los términos en que se ha descrito aquí esta posible puesta en práctica de la estrategia; sin que por ello tengan que negar la presencia de la realidad psicológica pertinente como se describió al inicio ni el interés filosófico de las ideas expresadas al decir «tuve que hacerlo» o «no podía hacer otra cosa» (véase Williams 1981b). Pero ahora lo que sostengo es más sencillo: aun aceptando los términos polémicos con que ha sido descrita, esta puesta en práctica no resuelve la cuestión del dominio de la moralidad. La elucidación propuesta de la noción de exigencia o demanda es, en su origen, perfectamente general, ya que no parece haber una buena razón para pensar que todas las exigencias son exigencias morales, para pensar que otros valores prácticos no puedan convertirse en exigencias. Pero siendo así, imaginemos la situación de un agente que se encuentra en una posición tal que *o bien* aumenta su sensibilidad a las exigencias morales *o bien* aumenta su sensibilidad a las exigencias de otro tipo. Nada de lo expuesto en estas páginas, incluidas las concesiones, le ha dado razón alguna para decidirse por la primera opción. (Las cosas podrían haber parecido diferentes si la referencia a la «sublimidad» de las exigencias morales hubiera implicado algo más que un mero fragmento de fenomenología, como en ciertas teorías kantianas; pero nunca he tenido ni siquiera un atisbo de lo que se supone que es ese «algo».)

En cierto sentido, resulta triste que esa estrategia se tope con dificultades, pues de otro modo habría podido suscitar una interesante propuesta acerca de la cuestión que resta de las dos que distinguimos al principio sobre el alcance interno de la moralidad. En analogía a una conocida maniobra usada en las discusiones filosóficas sobre el escepticismo en general, la propuesta suscitada habría sido ésta: que deberíamos buscar los aspectos morales de una situación cuando hay alguna razón específica para sospechar que se presentará en ella una exigencia o una demanda de moralidad. Parte del interés de dicha propuesta se habría derivado de la urgente necesidad de decir algo acerca de la cuestión restante —urgente a causa de la posible amenaza posterior de que la moralidad engulla dema-

siado, *demasiado tiempo*—. Cuando se aprecia la vasta área de la moralidad, y a la vez se acepta meramente que las consideraciones morales siempre son pertinentes para las cuestiones sobre lo que uno debería hacer, puede surgir la amenaza de que nos creamos invariablemente obligados a buscar las consideraciones morales potenciales dentro de cada situación. La propuesta suscitada nos habría dado más tiempo para nuestras propias vidas, y también para las de los demás. En lugar de sostener, quizá, que hay una exigencia moral constante y preponderante de buscar cualquier consideración moral pertinente, la propuesta habría indicado que hay una exigencia moral constante y preponderante de no pasar por alto ninguna razón específica para sospechar que se presentará alguna *otra* exigencia moral. Así pues, no habría sido ninguna prueba de inmoralidad responder a la posible amenaza posterior de que la moralidad engulla demasiado con la réplica de que la constante búsqueda de la dimensión moral sería una (excesiva) pérdida de tiempo.

No cabe duda de que el merecido prestigio de la ciencia se vería menoscabado si ésta fuera tratada de manera «cientificista», considerándola una fuente de respuestas para todas las cuestiones teóricas, científicas y no científicas por igual (véase Platts 1983b: 2-3). De modo similar, el merecido prestigio de la moralidad sin duda se vería menoscabado —se ha visto sin duda menoscabado, en mi opinión— al tratarla de una manera «moralista»: si se considera que es la fuente de respuestas para todas las cuestiones prácticas, tanto morales como no morales. Esta discusión de las cuestiones sobre el alcance interno de la moralidad ha sido intencionalmente problemática y general: con toda intención se ha puesto el énfasis en problemas generales que, según me parece, arrojan dudas considerables sobre el «moralismo» irreflexivo, o incluso sobre el reflexivo.

5.2 Discrepancias morales

Tiene que resultar ahora imposible dudar que las ideas y las prácticas morales de diferentes personas y diferentes pueblos se han diferenciado con frecuencia en enorme medida. Hasta el menos «relativista» de los filósofos morales, Kant, insistió en la necesidad de que sus discípulos se ocuparan de los detalles de esas diferencias; y desde la época de Kant, la conciencia de esas diferencias, y tal vez de esos detalles, se ha vuelto un lugar común. No obstante, antes de tratar de sacar alguna conclusión acerca de la naturaleza de la moralidad a partir del hecho acordado de las discrepancias morales, la prudencia filosófica sugiere que registremos la vaguedad del hecho acordado observando algunas de las discrepancias

entre las discrepancias. Esto es, nuestro interés inicial debe dirigirse al *tipo* de aparato conceptual que ha de usarse para la descripción de las discrepancias morales, y, de haber aquí alguna teoría, no tiene que ser, sin embargo, una teoría moral.

En una sección anterior de este capítulo vimos varios ejemplos de sistemas de pensamiento práctico que no eran sistemas de pensamiento moral: el del emperador Justiniano, el código homérico del respeto, el código de honor que se pone de manifiesto en los duelos, incluso el fenómeno más próximo a nosotros del machismo. Nada se dijo acerca de lo que distingue a esos sistemas de pensamiento práctico de los sistemas morales; sin embargo, en el actual nivel intuitivo, los ejemplos sirven para enfatizar el requisito de que, cuando se consideran presuntos casos de discrepancias morales, resulta esencial estar razonablemente seguro de que los sistemas de pensamiento que se comparan son todos, en verdad, sistemas de pensamiento *moral*. Los casos de confrontación entre un sistema de pensamiento moral, por un lado, y un sistema no moral de pensamiento práctico, por el otro, parecen ser muy diferentes. Sea cual fuere el contenido del pensamiento de que cada uno de esos dos sistemas es en verdad un sistema de pensamiento moral, permítaseme por el momento sólo estipular que todos los ejemplos que voy a mencionar han de comprenderse *como si* ese pensamiento fuera verdadero respecto de ellos (a menos, desde luego, que el contexto ponga de manifiesto que la estipulación no es válida).

Otro señalamiento preliminar: una diferencia en la práctica, por notable que sea, no necesariamente es manifestación de una diferencia en los valores (*ex hypothesi* morales). Los conceptos de valores morales dados, y los principios morales generales que contienen esos conceptos, son con frecuencia sumamente abstractos: su papel en la determinación de prácticas específicas está mediado, entre otras cosas, por las creencias de los agentes sobre muchos hechos empíricos. Por lo tanto, las diferencias en las prácticas no necesariamente tienen algo que ver con una diferencia en los valores: pueden simplemente surgir de las diferencias en las creencias acerca de los hechos positivos. El pasaje de los principios morales generales a las acciones concretas específicas suele ser de un tipo complejo: si en alguna ocasión perdemos de vista esa complejidad, nos veremos llevados a detectar discrepancias entre valores y principios morales donde no las hay.

5.3 Conceptos y concepciones

Vale la pena mencionar esos señalamientos preliminares sólo porque se los olvida con frecuencia: un discurso vago sobre «diferentes ideas morales» puede contribuir a dicho olvido. Pero esa misma vaguedad puede contribuir a otros olvidos más importantes. Quizá el más significativo y común sea el de la distinción entre las diferencias de conceptos morales y las diferencias de concepciones morales.

Permítaseme presentar la explicación de la distinción general entre conceptos y concepciones por medio de tres ejemplos que no se relacionan con la moral. El primero proviene de las ciencias naturales (véase Platts 1980a; Wiggins 1980: 77-86). Supongamos que dos científicos introducen la expresión «electrón» para referirse a la partícula subatómica responsable de tales y cuales efectos en una cámara de niebla. Cada científico procede entonces a construir una teoría que pretende explicar, de un modo científico y basado en leyes, el efecto observado. Sin embargo, las teorías así construidas difieren en lo que respecta a las leyes fundamentales que rigen a los electrones, es decir, difieren en las propiedades fundamentales que atribuyen a los electrones. Diré que tal caso es el de un *concepto compartido*, el de electrón, junto con *concepciones diferentes* acerca de lo que es un electrón. Es decir, existe en primer lugar un concepto compartido, el de electrón, porque la mejor interpretación que puede hacer uno de los científicos con respecto al uso que le da el otro al término «electrón» se obtiene comparándolo con su propio uso del mismo término: «Con *electrón* él se refiere a los electrones.» El término posee el mismo significado literal en el habla de cada uno de ellos; ambos hablan del mismo sujeto. Por eso resulta que sus diferentes creencias teóricas acerca de las propiedades de los electrones se encuentran, desde el punto de vista lógico, en conflicto inmediato. Y ello es posible a causa del presupuesto de un referente compartido anterior e independiente de la formación de las creencias teóricas acerca de las propiedades del referente compartido: esto es, la «fijación de la referencia» anterior y parcial del término «electrón». Pero obsérvese, en segundo lugar, que el hecho de que, no obstante, existan concepciones diferentes acerca de un electrón surge de la diferencia dentro de sus respectivas teorías acerca de las leyes fundamentales que rigen a los electrones. Ambos científicos tienen creencias diferentes sobre *qué es* un electrón, sobre la *naturaleza* de los electrones, sobre las *propiedades esenciales* de los electrones. La diferencia entre esas creencias modales es lo que constituye la diferencia de concepción de ambos científicos. Éstos también pueden discrepar con respecto

a otras creencias que atañen a meras contingencias relativas a los electrones: uno de ellos puede creer que hay más electrones en la atmósfera de Plutón que en la de Venus, el otro puede negarlo. Pero esta diferencia de creencias no constituirá una diferencia de concepción (aunque, en un caso dado, podría resultar que surgió de esta última); las creencias que pasan a integrar la concepción de un científico acerca de lo que es un electrón tienen un carácter enteramente distinto.

En este ejemplo tomado de las ciencias naturales, está claro cómo se puede resolver la diferencia de concepción: mediante el despliegue de los métodos científicos habituales que incluyen la observación, la experimentación y la construcción de teorías. Para nuestros propósitos actuales no necesitamos detenernos en la cuestión de si siempre tiene que haber una respuesta correcta a una pregunta como «¿qué es un electrón?»; basta con que haya muchas respuestas incorrectas. Lo que debemos enfatizar antes de abandonar este ejemplo es que lo que nos obliga *en estos casos* a tratar de responder la pregunta es su carácter unívoco en relación con las diferentes teorías científicas —el hecho de que en estos casos hay realmente un concepto compartido y una consecuente e inmediata incompatibilidad lógica de las creencias—. Los casos de diferencias en los conceptos producirían confrontaciones totalmente distintivas.

El segundo ejemplo que vamos a examinar no es mío: está tomado, con ligeras modificaciones, de un artículo sumamente importante de W.B. Gallie (1955-1956). Considérese un concepto como el de *los campeones* en relación, digamos, con un juego como el fútbol, en el que ese concepto no ha de entenderse en términos «formales» como los de «los campeones de la liga» o «los ganadores de la copa», sino más bien en términos equivalentes a los de *los mejores jugadores de fútbol*. Con esta interpretación, la cuestión de quiénes son los campeones puede provocar un debate en apariencia interminable y notablemente apasionado; y esto, de acuerdo con Gallie, debe explicarse en términos del *carácter esencialmente debatible* del concepto involucrado. Gallie ofrece las siete características siguientes de un concepto esencialmente debatible: (I) es «estimativo» de una actividad dada; (II) la actividad en cuestión posee un «carácter internamente complejo»; (III) hay ordenamientos rivales en lo tocante a la importancia de las partes o los rasgos componentes de la actividad; (IV) la actividad admite considerables modificaciones a la luz de circunstancias cambiantes y acaso imprevistas; (V) en las disputas, cada parte reconoce que su propio uso del concepto involucrado es impugnado por los de otras partes, y cada parte tiene por lo menos cierta comprensión de los «diferentes

criterios» a la luz de los cuales las otras partes afirman que aplican el concepto en cuestión; (VI) el concepto se deriva de un «modelo original» cuya autoridad es reconocida por todos los usuarios del concepto contrincantes, y (VII) es posible que las disputas acerca del «uso correcto» del concepto permitan que el logro del modelo original «se sostenga y/o se desarrolle de manera óptima».

Esas características se encuentran, en efecto, en el concepto de los campeones que acabamos de mencionar, y sin duda Gallie tiene razón al pensar que también se encuentran en otros conceptos más importantes (arte, democracia, justicia social, una vida cristiana). Pero el interés presente de los conceptos y las disputas caracterizados por Gallie reside en que proporcionan otro tipo de ejemplificación de casos de diferencia de concepción dentro de un marco de conceptos compartidos.

Los conceptos que estamos considerando son distintivamente evaluativos o «estimativos». ¿Cómo podría compartirse un concepto semejante dentro del contexto de las disputas sobre su aplicación? La respuesta teórica general con respecto a cuándo se comparte el concepto nos es ahora familiar: cuando la mejor teoría interpretativa de los disputantes revela que éstos hablan del mismo sujeto. Pero la exposición de Gallie nos permite agregar algo más: quizá, muy en particular, que lo que hace posible ese compartir el concepto es el acuerdo parcial previo sobre un uso del concepto —el acuerdo previo respecto del «modelo original», el acuerdo parcial previo respecto de algunos «modelos paradigmáticos» del concepto de que se trata—. (Obsérvese que Gallie escribía mucho antes de que se pusieran de moda ciertas teorías actuales de la referencia.) Y esa consideración que abre posibilidades se complementa con otras: la apreciación de los «diferentes criterios» utilizados por los demás, e incluso el propósito común de sostener y desarrollar el logro del modelo original.

Mientras más se aclara la posibilidad de compartir el concepto, más problemática puede parecer la cuestión que sigue: ¿cómo es posible que surjan diferentes concepciones sobre lo que es, por ejemplo, jugar bien al fútbol? También en este caso Gallie nos permite decir algo sobre la cuestión. Pese al acuerdo parcial previo sobre los «modelos paradigmáticos» de la actividad en cuestión, pueden surgir diferencias acerca de cómo proceder a partir de esos modelos (y acerca de qué es lo que exactamente está sucediendo en un proceder semejante). La complejidad interna de la actividad ejemplificada como paradigma, los «ordenamientos rivales» consecuentemente posibles de sus partes o rasgos componentes y el carácter «abierto» de la actividad en relación con las circunstancias cambiantes: estos elementos pueden dar origen, de un modo natural e inteligible, a

concepciones diferentes de lo que es una buena ejecución de la actividad (y por lo tanto a diferentes ideas sobre lo que convirtió a los modelos originales en *paradigmas*). El reconocimiento de lo anterior puede servir para explicar un hecho notable: que en muchos de estos debates, ambos adversarios pueden estar dentro de una y la misma persona. Y el reconocimiento y la comprensión de esto bien puede servir para dominar la propensión común a identificar cualquier adversario externo con el fatótum del diablo.

Este segundo ejemplo posee similitudes obvias —y diferencias igualmente obvias— con el primero que tomamos de las ciencias naturales. La similitud más obvia consiste en el papel clave que desempeña, para dar seguridad de que se comparten los conceptos, la «fijación de la referencia» previa y parcial de la expresión de que se trata, mientras que una diferencia obvia es la naturaleza evaluativa de los conceptos que intervienen en el segundo tipo de ejemplo. Sin embargo, surgen otras diferencias importantes si se considera el problema de cómo pueden resolverse las diferencias de concepción. En el primer tipo de ejemplo, resulta claro que el despliegue de los métodos de la ciencia natural constituye la base para resolver las diferencias de concepción. Ahora bien, en el segundo tipo de ejemplo, los resultados de tal despliegue pueden servir, en ocasiones, para socavar ciertas concepciones erróneas; y la capacidad de tal socavamiento se puede incrementar recurriendo además a los métodos de las ciencias sociales y a otros métodos similares de investigación. Las concepciones pueden descansar, de un modo complejo, en supuestos acerca de las cuestiones de hecho correspondientes. Pero incluso si tales supuestos acerca de esos hechos no contienen errores, las disputas en el segundo tipo de ejemplo pueden continuar razonable e inteligiblemente. Por lo general, cabe comprender tales disputas como disputas que giran, principalmente, en torno a la cuestión del ordenamiento, según su importancia —su pertinencia—, de las partes o rasgos componentes de la actividad en cuestión; y a menudo es posible entender que ésta es una cuestión que gira en torno a otras cuestiones, tanto sobre la mejor forma de adaptación de la actividad a circunstancias nuevas, como sobre lo que constituye exactamente un desarrollo positivo de esa actividad. Estas cuestiones se prestan tan poco a una resolución deductivamente razonada como a una decisión directamente científica; pero ello no demuestra que tales cuestiones representen un callejón sin salida, así como tampoco demuestra que las presuntas respuestas a los mismos no pueden ser más, o menos, razonables. Los intentos por cambiar la manera en que otro ve los modelos paradigmáticos de la actividad respectiva, al igual que los

intentos por hacerlo imaginar o apreciar vívidamente lo que sería la actividad si se realizara su concepción preferida de ella, no necesariamente son estudios de mera manipulación.

Aceptar que nociones como lo irrazonable, la miopía y la ceguera pueden hallarse en juego en estas disputas entre concepciones rivales de cierta actividad no significa todavía aceptar cualquier idea de *la única concepción verdadera*. Gallie parecía considerar que era por lo menos posiblemente deseable que tales disputas continuaran sin fin; pero su pensamiento presuponía la idea de una actividad «sostenida y/o desarrollada de manera óptima». Si hubiera que entender esa última idea en términos del límite de *la* «manera óptima», la posición de Gallie resultaría análoga a la que sostiene que, a la luz de hechos conocidos sobre la complejidad infinita del mundo natural y de nuestras propias y restringidas capacidades epistémicas, sería deseable una pluralidad de teorías científicas rivales sobre el mundo natural para alcanzar *la* teoría científica verdadera del mundo en relación con ese mundo. Pero el discurso de Gallie sobre la «manera óptima» no requiere ninguna posición semejante y, de todas formas, dudo que sea correcto atribuírsela. Consideremos un ejemplo pertinente, el de un equipo de fútbol. Es obvio que tenemos el concepto de un buen equipo de fútbol y, también, el de ser un equipo de fútbol mejor que otro. Por lo tanto, no es sorprendente que tengamos además el concepto del mejor equipo real. Pero, ¿poseemos un concepto del mejor equipo posible, o del equipo absolutamente perfecto? Desde luego que no: ¿cuál sería exactamente el *marcador*? Lo único que requiere lo que Gallie dice sobre la «manera óptima» es el concepto del ordenamiento comparativo aplicado a casos reales y a posibilidades específicas imaginables: no requiere ninguna idea del punto final absoluto de ese ordenamiento. Por lo menos en la gran mayoría de los ejemplos de este tipo carecemos de una idea semejante: si la idea de *la única concepción verdadera* de la actividad en cuestión está ligada a la idea de ese punto final, entonces tal concepción no existe. El reconocimiento de este hecho podría tener también un efecto saludable: el de servir para controlar la propensión a las hipérboles sin sentido durante este segundo tipo de polémicas entre concepciones rivales.

Pese a esas diferencias respecto de las formas y posibilidades de la resolución de las diferencias de concepción, la explicación final de concepciones en este segundo tipo de ejemplo no es tan diferente de la correspondiente al primer tipo de ejemplo relativo a la ciencia. Las creencias que determinan la concepción de una persona dada acerca de lo que es, por ejemplo, una vida cristiana, son ciertas creencias modales: sus

creencias sobre lo que es una vida cristiana, sobre la naturaleza de una vida cristiana y sobre lo que una vida cristiana exige. Tal vez sea verdad que en los ejemplos de este segundo tipo aparecerá un elemento mayor de relativización con respecto a circunstancias específicas dentro de los contenidos de estas creencias modales; pero ese elemento de relativización no implica una merma de la objetividad. Y es indudablemente cierto que otras diferencias de las creencias acerca de hechos —por ejemplo, si hay más personas que lleven vida cristiana en Sussex que en Yorkshire— pueden tener consecuencias sobre las diferencias de concepción: pero, al igual que antes, esto en nada socava el carácter distintivo de las creencias que constituyen la concepción de una persona dada de, por ejemplo, una vida cristiana.

El último ejemplo que consideraremos tampoco es mío: está tomado, con modificaciones considerables y tendenciosas, de la exposición de Philippa Foot del relativismo moral (Foot 1978a: 153). Consideremos ciertos juicios de gusto: juicios sobre quién es bien parecido, sobre qué comidas y bebidas saben bien, sobre qué colores combinan bien en los muebles o la ropa. Parece innegable que tales juicios pueden variar en enorme medida entre diferentes culturas, diferentes generaciones, diferentes clases y grupos sociales, e incluso entre diferentes individuos. Los antiguos mayas sostenían sin duda que Pacal era bien parecido, nosotros sostenemos lo mismo de Nureyev; los franceses tienen opiniones muy distintas de las de los habitantes de Nebraska sobre lo que es apetitoso; la cultura india parece más propensa a las combinaciones de colores fuertes que la de Surbiton. La descripción natural, prefilosófica, de cualquier caso semejante presupone que los juicios diversos se refieran a la misma cosa. Esto es, la descripción natural presupone que está en juego el mismo concepto a pesar de las grandes diferencias entre los juicios que lo utilizan. Pero, ¿cómo es posible compartir el concepto sobre un trasfondo de juicios radicalmente divergentes?

Uno de los contrastes con los dos primeros tipos de ejemplos salta de inmediato a la vista: no se cuenta ahora con la ayuda de la idea de una «fijación de la referencia» previa y parcial de la expresión en cuestión. Ocurre exactamente lo opuesto: gran parte de lo que explica las diferencias entre los juicios sobre quién es bien parecido es el hecho de las diferencias entre los paradigmas de lo bien parecido entre la cultura maya y la nuestra. Pero, entonces, ¿cómo puede estar en juego el mismo concepto?

Una vez más la respuesta teórica general a esa pregunta estará fundada en la teoría de la interpretación; pero, también una vez más, puede agregarse algo. Foot parece indicar que la clave ha de encontrarse en la generalidad de los conceptos en discusión: «[Los ejemplos más adecuados surgen] cuando no necesitamos fijar límites a las variaciones en la aplicación de una expresión o, más bien, no necesitamos fijarlos a su aplicación dentro de un dominio dado» (Foot 1978: 154). Aun así, podría parecer que esa generalidad lleva consigo una amenaza de vacuidad —como si la expresión «horrible» fuera el caso más favorecido—. En consecuencia, resulta importante advertir que en la discusión de Foot podría intervenir otra consideración que pudiera aducirse en este respecto. En un nivel intuitivo, los conceptos ilustrativos mencionados por Foot se entienden tal vez en términos disposicionales: de manera que un hombre bien parecido, por ejemplo, es quien tiende a producir una sensación agradable característica en quien lo experimenta (*tiene buena apariencia*). Y ahora la idea podría ser que lo que hace posible que el concepto se comparta pese a las grandes variaciones en su aplicación es el hecho de que se comparte la «respuesta sensitiva» producida por sus presuntos modelos.

Quizá también resulte natural decir que los antiguos mayas sustentaban creencias distintivas sobre lo que es ser bien parecido, sobre la naturaleza de la buena apariencia. Y entonces, a quien haya seguido esta discusión hasta este punto, le podría resultar natural decir que los mayas tenían una concepción distintiva de lo que es ser bien parecido. Pero antes de afirmar esto, deberíamos reflexionar sobre el hecho de que «ningún conjunto de [opiniones sobre quién es bien parecido] parece poseer más pretensiones de verdad que otro» (Foot 1978: 154). Aunque puede haber cierto grado de «objetividad local» en los juicios acerca de quién es bien parecido en tanto que se dan dentro de determinada comunidad o grupo que está de acuerdo en algunos paradigmas de lo bien parecido y en varios «criterios» para ser bien parecido, al considerar el caso de otra comunidad o de otro grupo que tenga opiniones radicalmente diferentes parece que no habría nada que decir. En realidad, la idea de que los antiguos mayas estaban equivocados o eran irrazonables, unilaterales o ciegos en sus juicios acerca de quién era bien parecido no parece ser sino la expresión de un imperialismo irreflexivo. Dado que deseo fijar la noción de una concepción en función de ciertas creencias modales distintivas —creencias, por ejemplo, acerca de la naturaleza, o la esencia, de la buena apariencia—, prefiero decir que la consecuencia de liberarnos del imperialismo irreflexivo es el abandono de toda concepción sobre lo que es ser bien parecido. Toda concepción será una mala concepción. Quizá

todavía tengamos que hablar de la «concepción» de nuestro imperialista acerca de la buena apariencia, pero hablar así ya significa señalar su error.

Lo anterior ha sido un intento por introducir, en términos generales, las distinciones —y algunas de las complejidades de las distinciones— entre la diferencia de conceptos, la diferencia de concepciones y la diferencia de «meras» creencias respecto de un algún hecho. Suponiendo que esas distinciones se aplican *de una u otra forma* dentro de la esfera moral, surge el resultado nada sorprendente de que una diferencia en la práctica moral puede ser la manifestación de diversos tipos de diferencias complementarias: de conceptos, de concepciones, de «meras» creencias y de las varias combinaciones entre estas cosas. Tampoco en este caso hay nada que deba asombrarnos. Conceptos como el de justicia —o democracia, o corrupción, o lealtad— son sumamente abstractos, al igual que los principios morales generales que contienen. El papel de tales conceptos y principios en la determinación de prácticas específicas está mediado por un buen número de elementos adicionales: las creencias del agente en cuestión acerca de muchos hechos, sus demás valores (sus otros conceptos evaluativos, morales o no morales), sus concepciones acerca de los valores pertinentes, sus creencias dentro de un contexto dado sobre las relaciones entre esos valores. Hablar simplemente de «diferentes ideas morales» sólo sirve para borrar la complejidad de esa mediación y para convertir la idea del reconocimiento de las discrepancias morales en una perogrullada inútil que enmascara las complejidades.

Es importante reconocer que cada una de las diferencias mencionadas —de conceptos, de concepciones, de «meras» creencias— es una diferencia de grado. Muchos filósofos de la ciencia se han concentrado en un supuesto límite para el caso de los conceptos científicos: se han concentrado, vale decir, en el caso supuestamente posible de dos teorías científicas que pretenden aplicarse al mismo, o casi al mismo, dominio de la experiencia, pero cuyos términos básicos pretenden referirse a entidades de naturaleza radicalmente diferente —es decir, cuyos conceptos básicos son radicalmente diferentes—. Esa supuesta posibilidad es aproximadamente lo que ha dado en conocerse como la *inconmensurabilidad radical* entre esquemas conceptuales científicos rivales: quienes practican esas teorías rivales viven, se nos informa, en mundos diferentes. No me interesa aquí si existe o no realmente tal posibilidad en el caso de la ciencia; todo lo que quiero *indicar* en este caso es una razón para pensar que a las diferencias entre los conceptos morales les falta mucho para llegar a esos supuestos límites. La idea es, en suma, la siguiente: si alguna vez nos decidimos a

usar la expresión «moralidad» de tal suerte que respete la singularidad de la moralidad, que respete, por ejemplo, el carácter distintivo del valor moral como sólo un tipo de valor práctico, entonces creo que nos veremos obligados a reconocer que cualquier esquema conceptual *moral* compartirá ciertos conceptos con el nuestro: los conceptos, por ejemplo, de justicia, de daño, de bienestar y, tal vez, de felicidad. Otras moralidades podrían incluir concepciones totalmente diferentes acerca de lo que son esas cosas; pero si ni siquiera se comparten esos conceptos, no puedo entender cómo puede considerarse que otros esquemas de pensamiento son *moralidades* rivales.

Dos observaciones finales para concluir esta sección. En primer lugar, resulta quizá muy poco claro, como cuestión de intuición lingüística, si hablar de «una moralidad diferente» se aplica a aquellos casos en que algún otro grupo evalúa los valores morales de diferente manera en comparación con otros valores no morales; y es igualmente poco claro si tal diferencia en la evaluación ha de contar siempre como una diferencia de concepción de los valores morales de que se trata (véanse las pp. 70–71). Pero lo importante no es disipar esas dudas principalmente terminológicas, sino más bien reconocer este fenómeno (en mi opinión muy difundido). Y en segundo lugar, aunque se ha supuesto aquí que las distinciones entre conceptos, concepciones y «meras» creencias se aplican de una u otra forma dentro de la esfera moral, nada se ha dicho aún sobre la(s) manera(s) precisa(s) en que se aplican. Introdujimos y refinamos esas distinciones mediante la consideración de tres tipos diferentes de ejemplos: pero nada hemos dicho sobre cuál es el tipo de ejemplo al que más se aproxima una discrepancia moral *específica*.

5.4 Cuestiones externas

Si llegamos a considerar que una moralidad es diferente de la nuestra, ¿cuál debe ser nuestra actitud? ¿Y cómo se relacionan las respuestas a esa pregunta con el tema de la objetividad del pensamiento moral? Estas cuestiones nos colocan firmemente dentro del terreno del tema del «relativismo moral»; pero el terreno donde así nos colocamos quizá no sea tan firme.

Por lo general, los filósofos morales no han salido muy bien librados de sus excursiones por este terreno. Tras comenzar subrayando el hecho de las discrepancias morales y combinan esto con alguna tesis acerca de la naturaleza subjetiva de la moralidad, luego han pretendido extraer una

conclusión «relativista» que imponga restricciones, usualmente considerables, a la posibilidad de la crítica moral de las sociedades que poseen moralidades diferentes de la propia. Casi todas esas tentativas de establecer una conclusión «relativista» han fallado de un modo patente y poco interesante. Una breve lista de las fallas más comunes puede servir para aclarar un poco el panorama.

(1) Se olvida el requisito de que el otro sistema de pensamiento que se considera sea, en sí mismo, un sistema de pensamiento moral.
(2) La conclusión que extrae el supuesto «relativista» es, en realidad, un principio moral «absoluto»; por ejemplo, en el sentido de que ninguna sociedad debería criticar, desde el punto de vista moral, a ninguna otra sociedad moralmente distinta (véase Williams 1972: 22-23).
(3) No se toman en cuenta las explicaciones obvias de las discrepancias morales en función de las diferentes circunstancias en las cuales viven otras personas —explicaciones que son a menudo compatibles con una concepción «no relativista» de la moralidad— (véase Platts 1979: 248-249).
(4) Se pasa por alto el hecho de que la relatividad de un predicado, por ejemplo, es compatible con la idea de que ese predicado posee condiciones «absolutas» de aplicación, una vez que se hacen explícitos los parámetros de su relatividad (piénsese en la relatividad, por ejemplo, de «es gramatical» como predicado de oraciones en relación con un lenguaje dado).
(5) Se da por supuesto que la categoría de aquello cuyo reconocimiento se determina mediante las necesidades, las preocupaciones y los intereses humanos coincide con la categoría de lo subjetivo (véase Platts 1983a: 147-148); de manera semejante, los dos usos que Hume hace indistintamente del concepto de «subjetividad» (véanse las pp. 114-116) no se distinguen (véase Wiggins 1987a: 201-202).
(6) Se supone que quien niega la restricción «relativista» impuesta a las posibilidades de crítica moral a una sociedad moralmente distinta tiene que condenar a los miembros de esa otra sociedad; de modo similar, se supone que la aceptación de una tesis «relativista» constituye la única forma razonable de fundamentar el valor de la tolerancia.
(7) Se supone que el reconocimiento de la naturaleza presuntamente subjetiva de la moralidad excluye cualquier posibilidad razonable de actuar de acuerdo con nuestra propia moralidad al establecer tra-

tos o compromisos con sociedades moralmente distintas (véase Foot 1978a: 161-162).

(8) Se desdeña la variedad de tipos de discrepancias morales que pueden producirse.

(9) Se pasa por alto el hecho de que normalmente se habla sobre las «diferentes concepciones de la misma cosa» dentro del contexto de la posibilidad de un debate razonable entre concepciones rivales.

La lista está lejos de ser exhaustiva, y todas las fallas especificadas tampoco son independientes unas de otras; no obstante, esa lista nos permite salvar una serie de escollos innecesarios. Al mismo tiempo, hay sin duda cuestiones genuinas acerca de, por así decirlo, el *alcance externo* de las moralidades. Podemos aproximarnos más a una apreciación de esas cuestiones si volvemos a concentrarnos en la distinción entre diferencias de concepto y diferencias de concepción, para considerar luego una dificultad general que se plantea para ciertos tipos de «relativismo», para algunas de las formas propuestas de restringir las posibilidades de la crítica moral entre diferentes moralidades.

Pensemos en dos grupos o sociedades que difieren en sus conceptos morales. En general, no tiene que haber nada de sorprendente en un caso así. Dadas, por ejemplo, las notables diferencias en las condiciones de vida de la burguesía de Madrid y las de los indios hopi de Norteamérica, lo sorprendente sería más bien que tuvieran exactamente los mismos conceptos morales. Pero aceptar la naturalidad y la inteligibilidad de esas diferencias en los conceptos no equivale a que se excluya la posibilidad de que uno de los grupos o sociedades critique razonablemente, desde un punto de vista moral, el marco conceptual moral del otro. En muchos debates morales que se dan en la práctica, el objeto de la discusión es exactamente si deben emplearse ciertos conceptos morales: los correspondientes a las virtudes y los vicios cristianos distintivos, los del patriotismo y el nacionalismo, quizá incluso conceptos sexistas y racistas. Es verdad que hay una cuestión teórica interesante acerca del tipo y el grado de comprensión de un concepto en discusión que puede alcanzar, en un debate así, alguien que rechaza el uso del concepto; pero, en muchos casos, ese rechazo no implica por fuerza una incomprensión absoluta. También es verdad que la naturaleza obcecada y a menudo ruin de un buen número de seres humanos no permite un gran optimismo acerca de las posibilidades reales de un cambio razonable dentro de su marco de conceptos morales: «El diablo es un optimista si cree que puede envilecer más a las personas.»

Pero la existencia de callejones sin salida no demuestra que viajar carezca de sentido (ni siquiera en la ciudad de México).

Hay una cuestión instructiva que surge para los casos en que existe una diferencia en los conceptos morales: ¿por qué no podrían combinarse de alguna manera los dos sistemas de conceptos morales? Una de las respuestas posibles consiste en que, precisamente por la naturaleza evaluativa y práctica de los conceptos de que se trata, el «esquema combinado» resultante podría resultar prácticamente contradictorio: podría resultar que el esquema combinado exigiera moralmente a un agente que ejecutara cierta acción y que a la vez le exigiera moralmente abstenerse de ejecutarla. Pero muchos han sostenido que tal contradicción práctica se suscitará también dentro de cualquier marco moral adecuado, dada la existencia de dilemas morales genuinos. De cualquier modo, hay otra manera de combinar dos sistemas de conceptos morales que no producirán inmediatamente esa supuesta dificultad: al estilo de quien trabaja por las mañanas como físico nuclear mientras que dedica sus tardes a estudiar las novelas de Jane Austen. ¿En qué yerra una persona que vive todos los días con el esquema moral de uno de los grupos hasta el mediodía, para luego cambiar al sistema del otro grupo?

Sin duda, muchas consideraciones diferentes podrían servir para determinar que una persona de hecho no critica moralmente tal combinación de sistemas diferentes de conceptos morales; y esas mismas consideraciones podrían servir para determinar que una persona de hecho no critica moralmente un marco de conceptos morales diferente del suyo. Aparte de la mera inercia, algunas de esas consideraciones son las siguientes:

(1) la persona cree que la cuestión no es lo suficientemente importante, digamos, porque cree o sospecha que el esquema del otro no hace ningún daño;
(2) la persona cree que la diferencia no tiene ninguna repercusión en cuanto a las posibilidades reales de adopción del otro esquema moral, como ocurre con esquemas de pensamiento moral históricamente remotos;
(3) la persona piensa que la crítica moral considerada no tendrá ningún efecto sobre el otro, sea esto debido a alguna creencia específica sobre el otro o debido a una creencia general en el sentido de que nadie presta atención a tales críticas;
(4) la persona teme ser o parecer moralmente pagada de sí misma;
(5) la persona teme comportarse de manera dictatorial, imperialista o colonialista;

(6) la persona es, en términos generales, tolerante con la diversidad, o incluso siente por ésta cierta atracción general;
(7) la persona no está segura de que haya algún error o incoherencia en el esquema moral del otro, sea esto debido a un escepticismo general acerca de sus propias capacidades «epistémicas» o debido a un escepticismo más específico relacionado con ese caso particular;
(8) la persona sospecha que no existe una diferencia moral sustancial entre su esquema moral y el del otro: quizá piensa que las diferencias equivalen meramente a una diferencia en los medios convencionales para alcanzar los fines en que concuerdan; o quizá cree que simplemente implican diferencias moralmente aceptables a la luz de los diferentes papeles que se desempeñan dentro de las distintas estructuras sociales;
(9) la persona cree que, en vista de otros rasgos no morales de la vida del otro, los juicios morales aislados acerca de él no vendrían al caso, resultarían de mal gusto o aun ridículos (tal vez juzga, por ejemplo, que el otro es un artista extraordinario, o incluso que se trata simplemente de un individuo notablemente ingenioso).

Esta mezcolanza de lista está inevitablemente incompleta. No obstante, pese al reconocimiento de la eficacia potencial de las consideraciones mencionadas para restringir las críticas morales de marcos morales distintos e, incluso, pese al reconocimiento de que esa restricción podría estar a menudo justificada, no debemos perder de vista la posibilidad, por lo menos aparente, de una crítica moral legítima de otros marcos morales —o del nuestro—; pues esa posibilidad parece ser un presupuesto de nuestro pensamiento acerca de los asuntos *morales*. Volvamos a considerar las tentativas de combinar diferentes sistemas de conceptos morales: seguramente forma parte de nuestras ideas acerca de lo que es el pensamiento moral el que rechacemos en esta área el tipo de «tolerancia» irrestricta que permitiría incondicionalmente todas esas combinaciones; y seguramente también forma parte de esas ideas el hecho de que no creamos que ese rechazo es una mera repulsión psicológica. Consideremos la combinación directa de dos esquemas diferentes de conceptos morales: creemos, me parece, que por lo menos en muchos casos los nuevos dilemas morales que se suscitarían en el esquema combinado serían ilusorios, no serían más que productos de un esquema equivocado; y por lo menos en muchos casos lo creemos, supongo, porque también creemos que por lo menos uno de los esquemas así combinados sería, de una manera u otra, erróneo. De igual modo, rechazamos la posibilidad completamente

general de un cambio sistemático al mediodía debido a que creemos que, por lo menos en muchos casos, uno de los esquemas combinados sería, en uno u otro sentido, erróneo. Pero si esto es correcto como descripción de nuestro pensamiento acerca de estos asuntos, entonces está claro que ese pensamiento presupone la posible aplicación de la noción de *error* a sistemas de conceptos morales, y ello a su vez presupone la posibilidad de la crítica entre sistemas diferentes de conceptos morales. Lo que esto muestra es el carácter por lo menos revisionista de cualquier negación completamente general de la posibilidad de la crítica moral significativa entre diferentes sistemas de conceptos morales. El mismo carácter por lo menos revisionista puede detectarse en la negación general comparable de la posibilidad de la crítica moral significativa entre partidarios de concepciones morales diferentes dentro de un marco de conceptos morales compartidos: *grosso modo*, tanto la persistencia de la discusión y el debate en algunos casos de este tipo como nuestra actitud, en algunos casos, hacia la posibilidad de un cambio a mediodía parecen indicar que existe la convicción de que, en muchos casos, está en cuestión algo más que la mera repulsión psicológica, y de que puede intentarse algo más que una mera manipulación psicológica.

Sin duda, a la posición «relativista» extrema todavía le queda algún espacio de maniobra, comenzando quizá con la adopción explícita del carácter por lo menos revisionista de las restricciones propuestas. Pero hay una estrategia más prometedora, y totalmente distinta, para quien esté interesado en descubrir alguna verdad en el «relativismo», a saber, la de buscar alguna restricción «relativista» más específica para las posibilidades de realizar críticas morales significativas entre moralidades diferentes. Si llegáramos a prestar la debida atención a distinciones importantes *dentro* de nuestro pensamiento moral, podría ahora ser la esperanza del «relativista», encontraríamos restricciones más cautelosas de las posibilidades de una crítica moral significativa, las cuales por lo menos no entran en conflicto con la estructura general de ese pensamiento y que hasta podrían ser exigidas por él.

5.5 Un relativismo contemporáneo

Esa estrategia ha sido explorada recientemente de diversas maneras. Gilbert Harman, por ejemplo, ha sostenido que se pone de manifiesto una importante restricción «relativista» a la posibilidad de la crítica moral significativa entre dos moralidades si distinguimos diferentes tipos de juicios morales. Los miembros de un grupo o una sociedad pueden decir

que algún miembro de otro grupo con prácticas morales diferentes es «(literalmente) un salvaje o [...] (literalmente) inhumano, perverso, un desleal, un traidor o un enemigo» (Harman 1975: 190); también pueden decir que, moralmente, los miembros de la otra sociedad no deberían hacer lo que hacen, que es una cosa terrible que actúen como actúan. Pero en lo que respecta a una acción cuya evaluación moral por parte de los miembros del primer grupo o sociedad se basa en alguna parte de su sistema moral no compartida por el otro grupo o sociedad, no tiene sentido que los miembros del primer grupo digan que algún miembro del otro grupo, desde un punto de vista moral, debería o no debería ejecutar la acción de que se trate, ni tampoco decir que estuvo moralmente bien, o mal, que la ejecutara. La razón de esa restricción, según Harman, estriba en que estos últimos juicios son «juicios internos» y una característica de tales juicios consiste en que sólo pueden expresarse acerca de una persona cuando se supone que ésta es capaz de ser motivada por las consideraciones pertinentes. En el tipo de caso especificado, la diferencia entre las moralidades de los grupos o sociedades muestra, *ex hypothesi*, que la suposición no se aplica a la acción en cuestión.

La restricción «relativista» de Harman supone dos tesis distintas:

(1) Hay una clase de juicios, los «internos», cuya expresión en relación con un auditorio dado sólo tiene sentido o propósito si se supone que el auditorio es capaz de ser motivado por las consideraciones pertinentes.
(2) Los juicios acerca de lo que una persona debería moralmente o no debería moralmente hacer, junto con los juicios acerca de si la persona actuó moralmente bien o moralmente mal al hacer lo que hizo, son «juicios internos».

¿Cuáles podrían ser los argumentos para sostener esas tesis? Quizá el más interesante, pertenezca o no a Harman, es el siguiente: consideremos *ciertos* juicios que dicen que una persona particular tiene una razón para hacer, o para no hacer, una determinada cosa. Si una persona tiene una razón para actuar de cualquier tipo, sin duda tiene que resultarle posible actuar de acuerdo con esa razón: si tal posibilidad no existe, ¿qué sentido tiene decir que *él tiene* esa razón? Ahora bien, el rasgo distintivo del *tipo* de juicios sobre razones que estamos considerando consiste en que la posibilidad de que el agente actúe de acuerdo con la razón en cuestión exige que esa razón esté conectada de un modo *específico* con el acervo actual de deseos del agente, con su «conjunto motivacional» actual: si *esa* conexión

no existe –si la razón en favor de ejecutar la acción de que se trata no se relaciona ni con su acervo actual de deseos ni con ningún acervo de deseos ampliado al que el agente pudiera verse razonablemente llevado a partir de su acervo actual mediante la eliminación de su ignorancia innecesaria acerca de «meros» hechos y la eliminación de las deficiencias innecesarias de su razonamiento lógico–, entonces carece de sentido decir que tiene una razón para actuar *de este tipo*. Tales juicios sobre razones son «juicios internos», y no veo ninguna razón para negar la existencia de una clase de juicios semejantes. (Por lo menos: no veo razón alguna para negarlo en tanto que no se suponga que esa clase de juicios –de *usos* de oraciones, si se quiere– puede identificarse, en general, haciendo referencia a alguna clase de las oraciones mismas.)

Si pudiera ahora demostrarse que los juicios sobre lo que una persona debería hacer moralmente, o sobre lo que sería moralmente correcto que hiciera, son siempre equivalentes a juicios acerca de razones *de ese tipo*, la restricción «relativista» de Harman quedaría establecida. ¿Cómo podría demostrarse esto? Un argumento relativamente extremo es el que sigue: está claro que estos juicios morales pretenden ofrecer razones para actuar, está claro que son juicios de razones; pero los *únicos* juicios de razones coherentes son los pertenecientes al tipo específico descrito en la primera parte de este argumento; de suerte que o bien esos juicios morales son incoherentes, o bien equivalen a «juicios internos»; puesto que la naturaleza, en general poco problemática, del empleo de tales juicios morales vuelve improbable la primera opción relativa a la incoherencia, puede concluirse razonablemente que tales juicios son en verdad «juicios internos». Otro argumento, en apariencia más moderado, es el siguiente: está claro que los juicios morales en cuestión pretenden dar razones para actuar, está claro que son juicios de razones; más aún, por lo menos a veces, los agentes actúan subsecuentemente conforme a esas razones; pero tales casos de acción conforme a esas razones sólo se comprenden de manera adecuada en términos de esa explicación específica del actuar conforme a cierta razón que se presupuso en la primera parte de este argumento; así que, de nuevo, los juicios morales en cuestión se revelan como «juicios internos». Precisar exactamente hasta qué punto es moderado este segundo argumento dependerá de las otras razones que han de ofrecerse para sostener la afirmación de que los casos en que se actúa conforme a estas razones morales sólo se comprenden de manera adecuada en términos de la explicación específica del actuar conforme a una razón que se hallaba presupuesta en la primera parte del argumento global: si se afirma que ésa es la única explicación coherente que puede darse del

actuar conforme a una razón, el aire de moderación desaparece. Pero incluso si se afirma, en términos más moderados, que esa explicación del actuar conforme a una razón constituye, simplemente como hecho, la explicación correcta de la acción conforme a razones morales, bien podrían suscitarse algunas dudas.

El primer aspecto que ha de observarse es ya un viejo conocido: el juicio escueto de que una persona debería hacer moralmente alguna cosa particular o de que sería moralmente correcto que lo hiciera —o incluso que posee una razón para hacerlo—, resulta notablemente carente de interés si es tanto el principio como el final del asunto (véanse las pp. 165-166). El interés y el valor de un juicio semejante cualquiera dependen de que, en potencia, estén respaldados por una especificación detallada de por qué debería hacer eso, de una especificación detallada de la razón para hacerlo —alguna especificación detallada, en términos de Harman, de *las consideraciones morales pertinentes*—. Pero supongamos ahora que, en un caso dado, se presenta en verdad un respaldo semejante: se produce cierta especificación detallada de la acción en cuestión como un medio para respaldar el aserto de que la persona en cuestión debería moralmente ejecutarla. Entonces el segundo aspecto que ha de observarse es que ahora no resulta nada obvio que la explicación emitida sobre cómo opera el juicio respaldado desde el punto de vista motivacional sea siempre correcta.

La explicación emitida es, en un claro sentido, una explicación humeana (modificada). La persona que se encuentra en el extremo de la recepción del juicio moral posee un acervo inicial de creencias y deseos (su «conjunto motivacional»). En el caso humeano más simple, el juicio respaldado tiene una conexión directa con ese acervo, de tal suerte que el agente, razonablemente, llegaría a tener el deseo de ejecutar la acción de que se trata. Pero el juicio respaldado podría tener otra conexión, esta vez indirecta, con ese acervo inicial al tener el mismo tipo de conexión directa con algún acervo hipotético modificado que se encuentre en una relación específica con el acervo actual de la persona: en esta relación, el acervo hipotético y modificado de creencias y deseos surgiría razonablemente del acervo actual de la persona si ésta se liberara del error innecesario en sus creencias acerca de «meros» hechos y de las deficiencias innecesarias de sus facultades de razonamiento (véase Williams 1981a). La sugerencia humeana (modificada) que se halla en juego consiste, pues, en que un juicio respaldado puede tener eficacia motivacional únicamente si está en una conexión apropiada, directa o indirecta, con el acervo actual de

creencias y deseos del agente —esto es, sólo si está en la conexión directa apropiada con el acervo actual de creencias y deseos del agente o con algún acervo hipotético modificado de tal índole—. Pero si los deseos actuales de la persona se diferencian en grado apreciable de los de quien emite el juicio moral, aquella condición bien podría no cumplirse; y entonces carecerá de sentido, según piensa el «relativista», la emisión del juicio moral respaldado acerca de lo que debería hacer la persona o lo que tiene una razón para hacer.

Pero, ¿exactamente por qué se considera que la explicación humeana es *la* correcta? A menudo se detecta en este punto un juego de prestidigitación. Quienes pretenden defender la aplicabilidad universal de esa teoría proceden con frecuencia criticando a la que es, quizá, la teoría rival más conocida: la teoría rival que se propone mostrar que la Razón requiere que una persona posea ciertos deseos, sea susceptible a ciertas consideraciones motivacionales, independientemente de cuáles pudieran ser sus deseos actuales. En términos más específicos, esta explicación rival pretende mostrar que una persona será en verdad susceptible a ciertas consideraciones motivacionales morales específicas, so pena de ser *irracional*. Pero los defensores de la explicación humeana (modificada) no tienen por qué encontrar dificultad alguna para mostrar que la tesis de la irracionalidad no es, en general, más que una baladronada; y concluyen entonces que su propia explicación alternativa del asunto es obviamente la correcta.

Pero ese resultado no se da tan fácilmente: Hume y Kant, o los humeanos y los kantianos, no agotan las opciones en este aspecto. Esto se puede advertir de varias maneras. Una preocupación acerca de la explicación humeana (modificada) es la siguiente: incluso en los casos que parecen ajustarse mejor a esa explicación —por ejemplo, un caso en que la persona que se encuentra en el extremo receptor del juicio moral emitido tiene el deseo general pertinente (por ejemplo, ser leal a su familia)—, se desdeña la cuestión de cómo ha de entenderse ese deseo y, por lo tanto, del tipo de deseo que es. El asunto no estriba, por lo menos en un nivel general, en saber si se trata de un deseo exigido por la Razón; la cuestión es más bien, por lo menos en un nivel general, saber si el deseo de la persona es razonable, si se entiende como una respuesta razonable a su concepción de los hechos en cuestión (véanse las pp. 146-150). Y ello apunta a un posible papel motivacional distintivo de los juicios respaldados que estamos considerando, un papel que no es ni kantiano ni humeano: quien emite el juicio no necesita suponer la presencia previa del deseo general pertinente ni en el acervo de deseos actual de la persona ni en el acervo hipotético

modificado; más bien, su objetivo al emitir el juicio *respaldado* es tratar de cambiar la concepción que tiene la persona de los hechos en cuestión, con la esperanza de que la concepción cambiada suscite razonablemente en la persona un deseo de ejecutar la acción de que se trata.

La primera parte de este ensayo incluyó un intento por estructurar y defender la psicología filosófica que requiere esta opción distintiva. Pero la idea de una opción semejante no tiene nada de nuevo, ya que es por lo menos tan vieja como lo sería Sócrates si todavía viviera, de suerte que no hay muchas excusas, en cuanto a originalidad presente, para el juego de prestidigitación que entrañan los intentos por defender alguna explicación humeana de todos estos asuntos mediante la mera refutación de la explicación kantiana. Una vez que la opción alternativa se reconoce como tal, surge la posibilidad de que algunos juicios respaldados de razones, aunque no sean «juicios internos» como se entiende esa noción en las teorías humeanas (modificadas), no queden por ello aislados de la posibilidad de que las personas que reciben los juicios actúen conforme a las razones dadas.

Lo anterior tiene importantes efectos sobre la restricción «relativista» de Harman a la posibilidad de la crítica moral significativa entre moralidades diferentes. La restricción puede expresarse ahora de la siguiente manera: en lo tocante a una acción cuya evaluación moral por parte de los miembros de un grupo depende de ciertas susceptibilidades propias a la motivación, no compartidas por los miembros del otro grupo, carece de sentido para los miembros del primer grupo hacer «juicios internos» acerca de uno o varios miembros del otro grupo en relación con la acción de que se trata. Ahora bien, a la luz de lo que se ha dicho aquí, hay varios elementos oscuros en la discusión de Harman. Uno de ellos, que resulta de inmediato pertinente, se vincula con la cuestión de cómo deben entenderse los enunciados que hablan de «susceptibilidad a la motivación», de ser «capaz de ser motivado por las consideraciones pertinentes». Mirándolo bien, parece que el pretendido modelo de Harman es un modelo humeano no modificado: esos enunciados han de entenderse en términos de conexiones directas con el acervo de deseos actual de la persona (acervo también de «metas» e «intenciones» y «hábitos» y «disposiciones»). No obstante, supongamos que los argumentos presentados posteriormente por Bernard Williams (en Williams 1981a) en favor del modelo humeano modificado son aceptables para Harman. Incluso con esta amplificación de las posibilidades de la crítica moral significativa, se pasa por alto una posibilidad adicional: que la emisión de un juicio moral respaldado en

razones sea un intento por cambiar la concepción que la persona tiene de los hechos en cuestión, con la esperanza de que la opinión cambiada suscite entonces razonablemente, en la persona que recibe el juicio, un deseo moral de ejecutar la acción de que se trata. Este argumento distintivo para emitir un juicio moral queda implícita e incorrectamente eliminado por la restricción «relativista». Y una vez liberados de una vinculación tan fácilmente comprehensiva con cierto modelo humeano, o incluso humeano modificado, de la motivación, no nos sentiremos inclinados, creo, a decir que el juicio formulado así no es «un juicio moral hecho y derecho» (Harman 1975: 195) ni a decir que, en este tipo de casos, existe la pretensión de que la persona ubicada en el extremo receptor del juicio «es susceptible a ciertas consideraciones morales en un esfuerzo por volverla a ella o a otras personas susceptibles a esas consideraciones» (*ibid.*: p. 193). (La última frase de Harman ofrece, de paso, un sólido apoyo para atribuirle una explicación humeana *no modificada* de la motivación. Obsérvese, además, que los aspectos en los que aquí insistimos van en contra de las propuestas de Harman sobre las formas lógicas literalmente relativistas y de su explicación convencional de la naturaleza de la moralidad. Pero nótese también que esa última frase de Harman —a causa del «o a otras personas»— podría impulsarnos a reconocer otro caso en el cual el propósito de hacer juicios respaldados en razones es el intento de cambiar la concepción de alguna otra persona sobre los hechos en cuestión, de alguna persona distinta de la que se ubica estrictamente en el extremo receptor del juicio. Piénsese, por ejemplo, en la formulación de un juicio respaldado semejante acerca de los *extranjeros*, en el que el público al que pretende dirigirse el juicio es cierta porción de la población nacional; pero adviértase también que, dentro de este contexto, la emisión de un juicio semejante puede tener muchos otros propósitos posibles —como la adulación de la población nacional, por ejemplo—.)

Debe darse aquí un reconocimiento explícito a un aspecto importante. Hay una buena pregunta acerca de cuándo resulta en verdad posible cambiar la concepción de una persona de los hechos en cuestión, y también una interesante pregunta subsidiaria acerca de cuándo existe una esperanza razonable de que el cambio de concepción produzca el correspondiente deseo de actuar. Aquí hay, ciertamente, un vasto terreno de problemas «relativistas»; sin duda, las respuestas a estas cuestiones harán referencia al carácter y a la personalidad de la persona en cuestión y, más específicamente, a sus deseos. Pero en este caso sólo tenemos que poner atención en dos cosas. En primer lugar, con respecto al asunto del «relativismo moral» *como aquí lo entendemos*, es esencial que la persona en

cuestión posea alguna moralidad (ni los marcianos de Harman ni el Hitler de Harman satisfacen claramente esta condición (1975: 191–193)). Y en segundo lugar, no hay razón para suponer que las referencias hechas a los deseos de la persona en las respuestas a dichas cuestiones sirvan para restablecer, como comprehensivamente correcta, una explicación de la motivación humeana o humeana modificada.

Dos comentarios finales sobre la provocativa defensa de Harman del «relativismo moral». Está claro que una diferencia de conceptos morales constituye, para Harman, una diferencia de susceptibilidades motivacionales: si una persona no tiene los conceptos que corresponden a las consideraciones morales pertinentes, difícilmente puede ser «susceptible» a la motivación por ellas (según la manera como Harman entiende «susceptibilidad»). Pero el asunto no es tan fácil en los casos en que hay diferencias de concepciones dentro de un marco de conceptos morales compartidos. Si Harman continúa con su explicación humeana original no modificada de la motivación, ese caso parece constituir una diferencia de susceptibilidades motivacionales. Pero si Harman aceptara el modelo humeano modificado de la motivación, el asunto no es tan simple. Es probable que la diferencia de concepción se vuelva manifiesta en una diferencia de acción y juicio y de deseo real; pero lo que cuenta en el modelo humeano modificado es cierto acervo hipotético de deseos modificado al cual la persona podría verse razonablemente conducida, dados sus deseos reales. Se plantea, entonces, la pregunta de si alguien que emplee este modelo humeano modificado reconocerá como ejemplo pertinente de tal modificación razonable de un acervo de deseos un caso en el que esa modificación se produce mediante una modificación razonable de la concepción real de la persona; si éste se reconoce como caso pertinente, entonces surge la otra pregunta de si quien emplea el modelo humeano modificado piensa que este proceso de modificación razonable de la concepción es potencialmente convergente con la única concepción verdadera. Si quien emplea el modelo humeano modificado pensara esto último, no pensaría que una diferencia de concepción real es una diferencia de susceptibilidades motivacionales, según entiende esta noción. Pero parece improbable, por decir lo menos, que algún humeano piense algo semejante.

El otro comentario final acerca de la discusión de Harman se relaciona con lo siguiente: él afirma que, en contraste con otros argumentos relativistas más tradicionales, su relativismo es neutral respecto de la cuestión de la objetividad o la subjetividad del pensamiento moral. Mucho de lo

que se ha dicho aquí en cuanto a los argumentos complementarios que acaso requiere la posición de Harman y en cuanto a su necesaria clarificación, parece indicar la presencia de tendencias subjetivistas en cualquier desarrollo y estructuración plenos de su relativismo. El último aspecto que hay que observar es que los medios obvios para tratar de arrojar dudas sobre la coherencia de los juicios de razones distintivos, de los *usos* distintivos de las oraciones acerca de razones, que aquí se han aducido contra la restricción relativista de Harman, así como sobre la coherencia de los juicios afines en el sentido de que una persona particular debería ser motivada moralmente por alguna consideración moral y debería moralmente tener el concepto de la consideración moral pertinente (véanse las pp. 118-119), presuponen todos alguna tesis subjetivista (revisionista). De suerte que, contra sus objetivos declarados, parece probable que el relativismo de Harman termine por apoyarse en un subjetivismo tan poco defendido como revisionista.

5.6 Otro relativismo contemporáneo

Otra exploración reciente de las posibilidades de un «relativismo moral» cauteloso se encuentra en los escritos de Bernard Williams (1974-1975; 1985: cap. 9). En el curso de esos escritos, las opiniones de Williams, en relación con diversos temas, fueron cambiando. Lo que sigue es, por lo tanto, una breve *reconstrucción* de una posición «relativista» sugerida en distintas partes de esos escritos y no una posición que pueda atribuirse claramente a Williams.

La médula de la posición «relativista» surge en los casos que ejemplifican la posibilidad siguiente en relación con dos grupos o sociedades S_1 y S_2 que poseen conjuntos de creencias morales «incompatibles»: debido a ciertos factores, como la industrialización, el avance del conocimiento o, incluso, la Marcha de la Historia, no existe ningún grupo —ni S_1 ni S_2 ni ningún otro grupo— que cuente con la posibilidad auténtica de adoptar las creencias morales de S_1 y vivir de acuerdo con ellas, y la posibilidad comparable de adoptar las creencias morales de S_2 y vivir de acuerdo con ellas, en el que las posibilidades negadas se entienden en términos tales que habrían tenido que haber sido compatibles por lo menos con un mínimo de sensatez —vale decir, con la conservación de cierto asidero en la realidad— y por lo menos con un mínimo de racionalidad (esta última condición compleja determina buena parte del contenido de la idea de una auténtica posibilidad, o imposibilidad, de adoptar un conjunto de creencias y vivir de acuerdo con ellas). En cuanto a una forma sensata y

racional de vivir cierto conjunto de creencias morales, no está en cuestión ninguna decisión práctica real en esa «confrontación» entre las dos sociedades: la pregunta «¿cuál de los conjuntos de creencias morales debe adoptarse para vivir conforme a él?» no constituye una cuestión práctica genuina para ningún grupo o sociedad. En palabras de Williams, la «confrontación» es «nocional», no «real» (Williams 1985: 160-161). Por ejemplo: debido quizá a cierta nostalgia infantil, los miembros de la sociedad mexicana contemporánea podrían creer que son capaces, aquí y ahora, de adoptar auténticamente las creencias morales de los aztecas precolombinos; pero a menos que esos mexicanos perdieran a la vez su cordura y su racionalidad, lo más que pueden hacer es remedar, de una manera grotesca, los *accesorios ornamentales* de la vida moral azteca; y lo mismo puede afirmarse con respecto a las posibilidades reales de que cualquier otro grupo o sociedad que exista en el presente adopte la moralidad azteca.

La facilidad de adoptar de un modo sensato y racional cierto conjunto de creencias morales es, cuando resulta posible, cuestión de grado; pero para el tipo de «relativismo moral» que ahora consideramos, la clave estriba en el punto de ruptura en que deja de existir tal posibilidad. Por qué se supone que esto es así se pone de manifiesto de la siguiente manera. En primer lugar, se nos invita a distinguir una clase de juicios morales muy generales acerca de la aceptabilidad moral de las creencias morales globales de una sociedad y, por consiguiente, acerca de las estructuras morales generales que se hallan dentro de esa sociedad. Así, por ejemplo, no están en cuestión los juicios sobre el comportamiento y el carácter de los individuos en la sociedad de que se trate. Luego, en un nivel intuitivo, el «relativista» nos pide que consideremos la cuestión de cuál podría ser el «propósito o la sustancia» (Williams 1974-1975: 142) de que los miembros de una sociedad formularan uno de esos juicios morales muy generales sobre la aceptabilidad moral de la moralidad de alguna otra sociedad en un caso en el que la relación entre las dos sociedades se parece a la existente entre S_1 y S_2, en un caso en que la confrontación entre sus moralidades es «nocional». Intuitivamente, sugiere el «relativista», en ese caso no tendría sentido, no tendría «propósito o sustancia», formular un juicio semejante.

Más aún, el «relativista» sostiene que es capaz de explicar y *justificar* esa intuición. La formulación de tal juicio moral muy general acerca de la moralidad de otra sociedad adquiere «propósito o sustancia» sólo dentro del contexto de una genuina cuestión práctica, sólo dentro de un contexto en que exista un grupo que afronte una opción auténtica entre las

dos moralidades en cuestión. ¿Por qué es ésa la única consideración que podría conferir «propósito o sustancia» a la formulación de esos juicios? Comparemos el caso, estructuralmente similar, en el cual los conjuntos de creencias en cuestión son científicos: por ejemplo, la teoría contemporánea de la combustión en contraposición con la teoría del flogisto. Y supongamos que, dados los requisitos de sensatez y racionalidad, junto con el avance del conocimiento científico y de la educación, no existe una posibilidad auténtica de que algún grupo o sociedad existente adopte la creencia en la teoría del flogisto. Ello no parece excluir la posibilidad de que los partidarios de la teoría contemporánea de la combustión formulen juicios generales significativos acerca de la inaceptabilidad científica de la teoría del flogisto. Los partidarios no se restringen a una evaluación científica de ciertos teóricos del flogisto; también pueden poner de relieve que la teoría del flogisto es falsa indicando sus incongruencias científicas específicas. ¿Por qué, entonces, no existe una posibilidad comparable en el caso de la moral, la cual pudiera conferir «propósito o sustancia» a la formulación de juicios morales de una generalidad comparable acerca de la aceptabilidad moral de algún otro sistema de creencias morales, incluso dentro del contexto de una confrontación «nocional»? Porque, afirma este «relativista moral», en el caso de la moral no existe la posibilidad de una crítica general *objetiva* de otros conjuntos de creencias morales: no hay manera de aplicar nociones de verdad o falsedad a esos juicios morales muy generales porque no se puede decir, hacer o investigar nada para conferirle sustancia a la aplicación de esas nociones. Podemos seguir llamando «falso», si queremos, al otro sistema de creencias morales; sin embargo, esa manera de expresarse sólo sirve para registrar el rechazo que sentimos por ese conjunto de creencias morales. Ésa es la (noción de) verdad en el «relativismo moral».

El «relativismo moral» que surge por esta vía es en verdad cauteloso: postula una restricción a la formulación de juicios morales pertenecientes sólo a una clase limitada, y sostiene que esa restricción entra en juego únicamente dentro de un contexto de confrontaciones de tipo específico. La tesis final del «relativista» es, pues, la siguiente: dada la falta de objetividad de los juicios generales acerca de la aceptabilidad moral de dos o más conjuntos de creencias morales, se sigue que, cuando el contexto de la confrontación es «nocional» respecto de dos de esos conjuntos, carece de «propósito o sustancia» la formulación, por parte de los partidarios de uno de los conjuntos, de juicios generales acerca del otro conjunto.

Es cierto que hay mucho más que decir para explicar y ejemplificar los términos en los que se expresa este «relativismo moral». También es cierto que existe la posibilidad de aumentar razonablemente el debilitamiento de la restricción propuesta para el juicio moral: esto podría lograrse mediante la consideración de otras fuentes prácticas del «propósito o sustancia» que pudieran surgir en los casos de confrontación «nocional», o bien mediante la consideración de otros casos en los que, pese a la «nocionalidad» y la subjetividad, hay suficiente relación, de algún tipo determinado, entre los dos conjuntos de creencias morales como para conferirle «propósito o sustancia» a la formulación de los juicios morales pertinentes (véanse MacIntyre 1981: cap. 17; Williams 1985: 165–167). Pero hay una dificultad más directa que se suscita con respecto a la restricción propuesta por este tipo de «relativista moral».

Admitamos la suposición de la falta de objetividad de los juicios preferidos por este «relativista». Si esto es así, ¿cómo puede haber algún «propósito o sustancia» en la formulación de un juicio semejante, incluso dentro de un contexto de confrontación que es real? El pensamiento del «relativista» es, supuestamente, que tales juicios, dentro de tal contexto, pueden tener «propósito o sustancia» en virtud de su influencia sobre las deliberaciones prácticas del grupo o la sociedad responsable de la «realidad» de la confrontación. Pero aquí tropezamos con el primo hermano de un viejo amigo: la mera afirmación de la «inaceptabilidad moral» de cierto conjunto de creencias morales carece tanto de interés como la mera afirmación de que una persona debería, moralmente, ejecutar una acción especificada (véase la p. 191). Cualquier aserción de uno u otro tipo sólo adquiere un interés sustancial si está potencialmente respaldada por juicios acerca de consideraciones más específicas. Pero el hecho de que este «relativista moral» niegue la objetividad en su clase preferida de juicios equivale a negar, por lo menos en términos generales, la posibilidad de algún respaldo potencial semejante para esos juicios. Resulta por lo tanto un misterio cómo puede suceder que, para este «relativista», sus juicios morales generales tengan «propósito o sustancia» incluso dentro de un contexto de confrontación de carácter «real».

5.7 ¿Perspectivas para un relativismo razonable?

Esto nos lleva al último de los exámenes contemporáneos de las posibilidades de un «relativismo moral» cauteloso que deseo considerar. En un artículo particularmente interesante (Foot 1978a), Philippa Foot, además de presentar una objeción contra Bernard Williams parecida a la

que acabo de plantear en contra de mi más hipotético «relativista moral», acomete dos tareas principales: en primer lugar, la de describir una estructura general que pueda discernirse en varias clases de juicios claramente «relativistas»; y en segundo lugar, la de explorar las dificultades que surgen cuando se intenta sostener que esa estructura relativista general está ejemplificada dentro del pensamiento y el discurso morales.

Para los fines de su primera tarea, Foot se concentra en los juicios de gusto: juicios acerca de quién es bien parecido, acerca de qué alimentos y bebidas saben bien, acerca de qué colores combinan bien en los muebles o la ropa. Conforme a la discusión que sostuvimos sobre tales juicios en las primeras páginas de este capítulo (pp. 180–181), podemos comenzar con la observación de dos de sus características pertinentes:

(1) esos juicios pueden variar mucho entre diferentes culturas, diferentes generaciones, diferentes clases y grupos sociales y diferentes individuos; y
(2) ningún conjunto de esos juicios posee más derecho a la verdad que otro: el grado de «objetividad local» que se encuentra en esos juicios no afecta ni remotamente la idea de que no hay nada que decir respecto de las comunidades o grupos con gustos radicalmente distintos.

Observamos también una preocupación acerca de la manera como los diferentes conjuntos de esos juicios pueden ser acerca de la misma cosa, pueden aducir el mismo concepto, dadas las afirmaciones (1) y (2). De suerte que, abundando en esa discusión anterior, podemos poner de manifiesto otros dos elementos en la estructura relativista de los gustos:

(3) los conceptos implicados en los juicios son «más bien generales» (Foot 1978a: 154); y, además,
(4) los conceptos implicados en los juicios se comprenden quizá informalmente en términos disposicionales: las disposiciones en cuestión son disposiciones a producir sensaciones y reacciones distintivas en quienes experimentan los objetos de tales juicios.

Hasta aquí por lo que se refiere a la estructura relativista general discernible en los ejemplos de Foot. ¿Qué ocurre ahora con la cuestión de si esa estructura se encuentra o no en el pensamiento y el discurso morales? La exposición de Foot se centra principalmente en el problema de si la afirmación (2) es verdadera cuando se aplica a juicios morales. Como ella observa, la idea que subyace en este tipo de «relativismo moral» nos es familiar: se trata de la idea de que en cualquier caso en que se presente una diferencia de opiniones morales, la discusión entre las partes no tardará

en derrumbarse, no tardará en estancarse, de suerte que no hay entonces un motivo razonable para persistir en la formulación de juicios morales acerca de la otra parte. Al respecto, Foot desarrolla dos argumentos. En primer lugar, el «relativista moral» se limita a afirmar que ahí termina el asunto, sin dignarse a considerar los detalles de las discusiones dentro de esos contextos de diferencias morales con el fin de fundamentar su aserto. Una vez que hacemos abstracción del cansancio, el aburrimiento y la impaciencia, no resulta claro *a priori* ni dónde ni cómo finalizan esas discusiones. La conclusión que Foot quiere sacar es que, les agrade o no, los filósofos morales tienen que prestar mucha más atención a los ejemplos de discusiones morales que se encuentran en la historia y en la literatura. Y en segundo lugar, el problema con la idea del «relativista» no estriba únicamente en su carácter *a priori*: si acaso descendiéramos a considerar los detalles de las discusiones reales dentro de los contextos de las diferencias morales, nos toparíamos con la dificultad que nos impone nuestra limitada comprensión de muchos de los conceptos involucrados en esas discusiones. Foot menciona como ejemplos los conceptos de *valor* y de *felicidad*: en los términos que aquí utilizo, sin contar con concepciones más adecuadas de esas cosas, no podremos alcanzar la comprensión necesaria de las discusiones en las que figuran.

Así pues, la exposición de Foot de este «relativismo moral» termina con una nota manifiestamente agnóstica: antes de que podamos llegar a alguna conclusión razonable sobre la verdad o la falsedad de tal «relativismo», tendremos que haber prestado más atención a los detalles de la discusión dentro de los contextos de diferencias morales y haber reflexionado más acerca de la naturaleza de algunos conceptos clave que figuran frecuentemente en esa discusión. Pero aunque las observaciones de Foot me parecen sustancialmente correctas, también me parece probable que las dificultades a las que se enfrenta este tipo de «relativismo moral» son aún más considerables.

En primer lugar, aunque sólo sea para rechazarla, vale la pena notar la maniobra que intenta establecer que no se puede decir nada razonable dentro de los contextos de diferencias de opinión morales mediante la demostración de que, en general, no existe la posibilidad de que una de las partes demuestre que la otra es *irracional*. Dada la falta, por lo menos en general, de un procedimiento semejante basado en la Razón Pura Universal como medio para resolver las diferencias de creencias morales, este «relativista» pretende sacar la conclusión de que, en general, no existe ningún procedimiento para resolver esas diferencias, y que, en general,

no puede decirse nada razonable con miras a una resolución semejante. Pero la dificultad a la que se enfrenta esa pretensión resulta ahora clara: del hecho de que no haya ningún procedimiento basado en la Razón Pura Universal para resolver las diferencias morales, no se sigue que no pueda decirse nada razonable con miras a ese fin. Kant y Hume no agotan más las posibilidades en este aspecto que en el de la teoría general de la motivación. Hay muchas cosas distintas que podrían expresarse razonablemente dentro de diferentes contextos de diferencias morales. Es probable que el asunto sea diferente en los casos en que hay una diferencia parcial de conceptos morales de como es en los casos en que hay una diferencia de concepciones; y es probable que el asunto sea distintivamente diferente cuando lo que está en discusión es el equilibrio entre las consideraciones morales y otras consideraciones evaluativas. Pero no debe sorprender que quiera yo enfatizar una posibilidad específica con respecto a lo que razonablemente puede decirse en algunas discusiones: se trata de la posibilidad de producir un juicio moral respaldado con la intención de tratar de producir un cambio razonable en el conjunto de creencias morales de la otra parte, mediante la producción de un cambio razonable de su punto de vista acerca de los hechos. Que esta posibilidad se presente en un caso dado dependerá, sin duda, del carácter de la otra parte y de sus creencias morales; de suerte que tenemos aquí otro terreno extenso y problemático cuyo mapa es preciso trazar antes de que podamos sacar conclusiones razonables acerca de las restricciones «relativistas» a las posibilidades de la crítica moral significativa entre moralidades diferentes.

El segundo aspecto que debe observarse se relaciona con una dificultad distintiva que enfrenta el tipo de «relativismo moral» que estamos considerando, una dificultad relacionada con los elementos (3) y (4) de la estructura relativista general. ¿Cuáles son los conceptos morales que, según se supone, son capaces de trascender las diferencias radicales de los juicios que los entrañan, aun cuando no se pueda decir nada con miras a una resolución razonable de esas diferencias? La condición de generalidad apreciable parece descartar la candidatura, por ejemplo, de los conceptos de vicios y virtudes específicos. Quizás los conceptos que presentan más probabilidades de satisfacer esa condición sean los de *la mejor acción desde un punto de vista moral, lo que moralmente debería hacerse, el deber moral* y, tal vez, *lo moralmente bueno*, o los que, como la *aceptabilidad moral* de Williams, se utilizan en juicios morales muy generales acerca de las moralidades, acerca de conjuntos globales de creencias y prácticas morales. En cualquier caso, el «relativismo moral» resultante enfrenta dos

nuevas dificultades: la de volver convincente la afirmación de que tales conceptos son de naturaleza disposicional de una forma bastante similar a la supuesta en (4), el elemento final de la estructura relativista general; y la de defender el interés de la restricción «relativista» resultante a la luz del escaso interés de los juicios morales expresados en términos de esos conceptos, cuando esos juicios carecen de algún respaldo en juicios acerca de consideraciones más específicas. Ese aspecto nos ha proporcionado razones que a estas alturas nos aburren tanto que ya no tendremos mucho interés en hacer esos aburridos juicios. Finalmente, debemos observar una dificultad que enfrenta cualquier «relativismo moral» que opta por conceptos como *la mejor acción desde un punto de vista moral, lo que moralmente debería hacerse* y *el deber moral*. ¿Qué distingue a ese «relativismo» de la posición de quien rechaza (correctamente) una concepción de la moralidad como «cálculo de acción»? ¿O de la de quien insiste (correctamente) en la dificultad general de determinar con exactitud lo que la moralidad exige en lo que a la acción se refiere? Aún está lejos de resultar evidente cualquier restricción «relativista» *distintiva* del juicio moral.

Quizá, como con tanta frecuencia ocurre en filosofía, lo importante es el camino, no el punto de llegada. No obstante, pese a la sutileza de estas últimas exploraciones de las posibilidades de una restricción relativista «cautelosa» para la crítica moral significativa entre moralidades diferentes, no está en modo alguno claro que se haya avanzado mucho hacia una formulación defendible de una restricción semejante. Desde luego, durante nuestra consideración de estas últimas exploraciones han surgido buenas cuestiones para una investigación ulterior; pero esas mismas cuestiones se habrían podido plantear, y tal vez con mayor claridad, sin los «marcos teóricos» de los «relativismos morales».

Quizá, también como con tanta frecuencia ocurre en filosofía, el asunto de la teoría sea engañoso. Esto es, quizá en este caso una actitud «relativista moral» vaga y mal definida haya llegado a adquirir el atractivo que tiene gracias a la fuerza negativa de sus críticas —a veces morales, a veces filosóficas, a veces justificadas y a veces no— de supuestos errores dentro de otras doctrinas de diferentes grados de «teoricidad». De modo que quizá, con el propósito de salvar «la verdad del relativismo moral», lo más que puede hacerse con provecho es la investigación poco sistemática y poco teórica de las diversas consideraciones que de hecho pueden hacer que la formulación de juicios morales entre moralidades diferentes resulte una actividad un poco absurda, o incluso de mal gusto. Existen tantas cuestiones *prácticas* genuinas e importantes acerca del alcance externo de

la moralidad como acerca de su alcance interno; y hay importantes conexiones entre los dos tipos de cuestiones. Así, por ejemplo, tal vez sea necesario sopesar los juicios morales sobre las moralidades de otras sociedades comparándolos con otros tipos de juicios sobre otros aspectos de esas sociedades. Cabe suponer que las cuestiones internas que a nosotros nos resultan problemáticas tienen sus equivalentes para los miembros de otra sociedad; y si éstos resuelven sus cuestiones mediante formas o balanzas diferentes –si tienen, digamos, opiniones distintas sobre el dominio de la moralidad–, no por ello ha de estar claro que haya mucho interés en los juicios morales aislados sobre la manera en que llegan a una resolución. Pero una vez que el filósofo moral ha reconocido y aclarado las cuestiones y las conexiones, quizá haya hecho todo lo que razonablemente cabía esperar de él. Quizá, por ejemplo, las cuestiones externas no se prestan a respuestas (restrictivas) fundadas en una teorización reflexiva más de lo que se prestan a obtener respuestas (expansivas) fundadas en una moralización no reflexiva.

Entre otras cosas, en este capítulo he tratado de mostrar las considerables dificultades que enfrenta cualquier intento de fundar una restricción «relativista» de las posibilidades de una crítica moral significativa entre sociedades moralmente distintas en el presunto carácter subjetivo de la moralidad; cualquier intento por deducir ese carácter a partir del presunto hecho del acuerdo sobre alguna restricción semejante se enfrenta a dificultades comparables. Al margen del interés intrínseco de estos intentos, había una razón particular para considerarlos aquí. En la primera parte de este ensayo se desarrolló una teoría del valor y la valoración que incluía un tratamiento cognoscitivo y no subjetivo de los miembros de cierta clase de valoraciones, y se sugirió entonces que, desde un punto de vista intuitivo, por lo menos algunas valoraciones morales eran miembros de esa clase; a continuación se sugirió que (a) algunas de esas valoraciones morales se comprenden como respuestas cognoscitivas a cosas meramente dignas de deseo, a cosas que son meramente objetos razonables de deseo, mientras que (b) otras de esas valoraciones morales se comprenden en los términos, más fuertes, de respuestas cognoscitivas a cosas que merecen ser deseadas, a las que les corresponde ser deseadas, que deberían ser deseadas (pp. 120-122). Si esas ideas intuitivas son correctas, se siguen dos cosas. En primer lugar, la presencia de las valoraciones morales mencionadas en (a) trae consigo, en ciertas circunstancias, la posibilidad de una crítica moral de sociedades moralmente distintas, de sus miembros y sus acciones: la crítica moral que puede surgir cuando se estima que, en

una sociedad moralmente distinta, ocurren valoraciones morales de este tipo, que están dirigidas a cosas que *no* son dignas de deseo. En segundo lugar, y más en particular, la presencia de las valoraciones morales mencionadas en (b) trae consigo, en ciertas circunstancias mucho más extendidas, otra posibilidad de crítica moral entre sociedades moralmente distintas, de sus miembros y sus acciones: la crítica moral que puede surgir cuando se considera que, en una sociedad moralmente distinta, hay una *diferencia* en las cosas hacia las que se dirigen ciertas valoraciones morales de este tipo. En cualquier caso, cuando se produce la posibilidad de una crítica moral, nadie tiene por qué negar que alguna de entre varias consideraciones pudiera servir para obstaculizar la formulación efectiva de la crítica (pp. 185–186); pero de acuerdo con las ideas citadas, la posibilidad de esa crítica moral continúa en pie. Es esa posibilidad la que muchos «relativistas» han querido negar.

La posición «relativista» extrema niega esa posibilidad a través de la negación completamente general de la posibilidad de las críticas morales significativas entre diferentes sistemas de pensamiento moral (pp. 187–188). No obstante, esa negación completamente general posee un carácter tan revisionista que su aceptación amenazaría nuestra comprensión de la idea misma de que los sistemas de que se trata son sistemas de pensamiento *moral* diferentes: la posibilidad de una crítica moral entre sistemas de pensamiento diferentes forma una parte constitutiva del hecho de que sean sistemas de pensamiento moral diferentes. Las posturas «relativistas» más cautelosas no tienen ese carácter arrollador, o al menos revisionista; en realidad, algunas pretenden incluso deducir una restricción «relativista» a partir de los términos de nuestro pensamiento moral mismo. Uno de esos «relativismos», el de Harman, niega la posibilidad de formular por lo menos algunas de las críticas morales autorizadas por los pensamientos intuitivos citados, a saber, ciertas críticas dirigidas a acciones particulares de miembros individuales de la sociedad moralmente distinta, en la que esas acciones emanan de las valoraciones morales distintivas de los miembros (pp. 188–189). Pero el argumento de Harman para negar esto, igual que un conocido argumento en favor de la subjetividad de la moral (pp. 201–202), pasa simplemente por alto la posibilidad del tipo de tratamiento cognoscitivo y no subjetivo dado a ciertas valoraciones morales de acuerdo con los pensamientos intuitivos mencionados. Más aún, en la medida en que se considere que los tipos de posibilidades de crítica moral aducidos contra la negación de Harman tienen que hallarse siempre potencialmente en juego entre sociedades moralmente distintas, en esa misma medida se considerará también que los pensamientos intui-

tivos mencionados tienen siempre que tener aplicación a los sistemas de pensamiento moral.

En este capítulo he tratado así mismo de ilustrar la «singularidad» de la moralidad mediante la presentación de ejemplos de valores prácticos que no son valores morales (pp. 167–168) y mediante una exploración tentativa de las consecuencias de esa «singularidad» para las cuestiones acerca del alcance interno de la moralidad en las vidas humanas normales. En un plano más teórico, en el de la pretendida identificación de los rasgos centrales de la teoría interna del pensamiento y la práctica morales, se hicieron varias propuestas:

(1) Los vocabularios primarios de la moralidad corresponden a vicios y virtudes específicos y a evaluaciones todavía más mundanas; por lo tanto, el área de la moralidad —la parte de nuestras vidas que potencialmente posee una dimensión moral— es casi «la totalidad de nuestro modo de vivir y la calidad de nuestras relaciones con el mundo» (p. 165).
(2) Cualquier sistema de pensamiento moral empleará los conceptos de justicia, de daño, de bienestar y, tal vez, de felicidad; puede haber concepciones variadas sobre lo que son estas cosas, pero los conceptos tienen que estar presentes para que un sistema de pensamiento sea un sistema de pensamiento moral (pp. 182–183).
(3) La posibilidad de la crítica moral entre moralidades diferentes forma una parte constitutiva del pensamiento moral mismo (pp. 186–188, 205).
(4) Esa posibilidad de una crítica moral entre moralidades diferentes está a menudo configurada por la presencia, en esas moralidades, de valoraciones morales entendidas en términos cognoscitivos y no subjetivos, y en ocasiones está configurada más exactamente por la presencia de valoraciones morales entendidas como respuestas cognoscitivas a cosas que ameritan ser deseadas, a las que les corresponde ser deseadas; quizá ocurra incluso que toda moralidad tenga que incluir valoraciones morales de estos tipos y, por lo tanto, tenga que autorizar la posibilidad de las formas correspondientes de crítica moral (pp. 204–206).

El reconocimiento de que (1) y (4) son verdaderos trae consigo la apreciación de la importancia que tiene, para la comprensión filosófica de la moralidad, la atención a los papeles que potencialmente desempeñan las razones específicas en el respaldo de juicios morales más abstractos, sean

esas razones específicas lo que antes se llamó «caracterizaciones de deseabilidad» o lo que ahí se llamó «razones» (p. 68; véanse las pp. 165–166). Las consecuencias de no prestar esa atención, de un tratamiento completamente abstracto de juicios ciertamente obtenidos por abstraccción, se pusieron de manifiesto en varias ocasiones durante la discusión desarrollada en este capítulo sobre los «relativismos morales». No es ninguna casualidad que este ensayo no se titule *Lo correcto y lo bueno*.

En el último capítulo examino los puntos de vista de dos filósofos que, según se ha considerado, han criticado a la moralidad misma. Procedo así por tres motivos. En primer lugar, creo que el examen de esos puntos de vista nos puede ayudar a apreciar *lo que no es la moralidad*: nos puede ayudar a apreciar que ciertas afirmaciones tendenciosas no forman parte de la teoría interna del pensamiento y la práctica morales. En segundo lugar, pienso que ese examen puede llevarnos a apreciar algunos rasgos desconcertantes del *uso* efectivo que hacemos del pensamiento moral, rasgos relativamente olvidados en las discusiones filosóficas contemporáneas sobre la moralidad. Y, por último, ese examen puede llevarnos también a apreciar un problema acerca del carácter de la motivación moral y el sitio que ocupa dentro de nuestro sistema conceptual moral. Los dos últimos puntos deben servir para eliminar cualquier impresión que pudiera surgir de que pienso que todo está bien con nuestro pensamiento y nuestra práctica morales.

Capítulo 6

CRÍTICOS DE LA MORALIDAD

> Los moralistas se han empeñado en derrotar al vicio y depurar el corazón de todos los apetitos e inclinaciones perjudiciales: Estamos en deuda con ellos por esto de la misma manera en que lo estamos con quienes destruyen a las sabandijas, y limpian a los países de todas las criaturas nocivas. Pero, ¿acaso no puede un naturalista practicar la disección a un topo, hacer experimentos con él e investigar la naturaleza de su hechura, sin ofender a los cazadores de topos, cuyo trabajo consiste exclusivamente en matarlos tan rápido como les sea posible?
>
> Bernard de Mandeville

> Al utilizar el cuchillo para practicar una vivisección en el pecho de las *virtudes mismas de su tiempo*, [los filósofos] traicionaron lo que constituía su propio secreto: conocer una nueva grandeza del hombre, un camino nuevo y sin hollar hacia su perfeccionamiento. En cada ocasión, pusieron al descubierto cuánta hipocresía, cuánta comodidad, que le permite al hombre liberarse y abandonarse, cuántas mentiras se enconden bajo las formas más honradas de su moralidad contemporánea, cuánta virtud quedó *rebasada*.
>
> Friedrich Nietzsche

6.1 *Terra incognita*

Así como Hume tuvo a su William Warburton, Mandeville tuvo a su William Law, según el cual el autor de *Fábula de las abejas* no era sino un hombre «que llegó como misionero del reino de la oscuridad para hacernos daño». (La peor de las suertes le tocó a Nietzsche, quien tuvo a su hermana; lástima que no se llamara Wilhelmina.) Y así como Hume es el gran ironista de la filosofía, Mandeville es uno de sus más grandes satíricos. Pero su estilo satírico no debe llevarnos a considerarlo una figura superficial, un mero guasón que se complace –y nos complace– en hacer desfilar juguetonas paradojas. Desde el punto de vista histórico, Mandeville tiende un importante puente entre las doctrinas de Hobbes y las de Hume; de mayor importancia para nuestros propósitos presentes

es que es posible interpretarlo como un defensor de tesis acerca de la moralidad que, de ser verdaderas, son importantes, y que, aun si fueran falsas, poseen un interés considerable.

¿Es el hombre naturalmente bueno? Y ¿cómo es posible la sociedad humana? Estas preguntas fueron dos de los pivotes sobre los cuales giró el debate entre los moralistas británicos de los siglos XVII y XVIII. En ese debate se introdujo el holandés Mandeville. Su descripción del origen de la sociedad y de los afectos sociales comenzó, como el de Hobbes, a partir de una concepción del hombre «natural», que posee el instinto innato de amarse a sí mismo y a nadie más:

> El hombre lo centra todo en sí mismo, y no ama ni odia sino por mor de sí mismo. Cada individuo es un pequeño mundo en sí mismo, y todas las criaturas, hasta donde su entendimiento y sus habilidades se lo permiten, se empeñan en hacer feliz a ese yo: ésta es su permanente labor, y parece constituir el designio entero de la vida. (Mandeville 1924: 178)

Ese instinto innato de amor propio se pone de manifiesto en un gran número de fenómenos: por ejemplo, la necesidad de autoconservación, el enojo, la envidia, el amor por la tranquilidad y el temor. (En realidad, como dice ese «gran teólogo» La Rochefoucauld, «por muchos que sean los descubrimientos que se hayan hecho en el mundo del amor propio, queda todavía mucha *terra incógnita* por descubrir» (Mandeville 1970: 240 [1982: 150]).) Pero Mandeville también incluye dentro de su concepción del hombre «natural» un instinto innato de *autoestima* completamente diferente:

> [T]odos nacemos con una pasión manifiestamente distinta del amor propio, la cual, cuando es moderada y se halla bien controlada, provoca en nosotros el amor por el elogio y el deseo de que los otros nos aplaudan y piensen bien de nosotros, y nos incita a las buenas acciones: pero esa misma pasión, cuando es excesiva o está mal encaminada, independientemente de lo que provoque en nosotros mismos, ofende a los otros, nos vuelve odiosos y se denomina orgullo. (Mandeville 1732b: 6–7)

> Todos los hombres son parciales en sus juicios, cuando se comparan con los otros; no hay dos iguales que piensen tan bien uno del otro como cada uno piensa de sí mismo [...] (Mandeville 1924: 271)

Al no existir previamente «ninguna palabra o expresión que comprenda la totalidad de los diferentes efectos de esta misma causa», Mandeville se

ve impelido a acuñar para esa causa «la palabra autoestima», que evidentemente ha de entenderse como un «término técnico». El amor propio y la autoestima son instintos innatos distintos, así como una cosa es luchar principalmente por mantenerse *a sí mismo* en la existencia y otra juzgar que *ese* ser por cuya existencia principalmente se lucha es el ser de más valor. Más aún, por lo menos en ocasiones, el gusto por sí mismo puede resultar un instinto más fuerte que el amor por sí mismo, como en el caso de quien comete suicidio para preservar su honor: así, el suicidio de Lucrecia fue «prueba indudable de que apreciaba menos su virtud que su gloria, y su vida menos que las otras dos» (Mandeville 1970: 223 [1982: 136]).

> El verdadero objeto del orgullo o de la vanagloria es la opinión de los otros; y el supremo deseo que un hombre poseído por ellos y enteramente lleno de ellos puede alentar consiste en que el mundo entero piense bien de él, que lo aplauda y lo admire, no sólo en el presente, sino en todas las épocas futuras. (Mandeville 1924: 64)

Y así, la tesis más general de Mandeville es la siguiente: semejante comportamiento humano que no puede explicarse desde el punto de vista de las operaciones del instinto innato del amor por sí mismo, puede empero explicarse de acuerdo con las operaciones adicionales del instinto innato diferente, y a veces más poderoso, del gusto por sí mismo.

Un aspecto crucial del comportamiento en cuestión consiste en la formación de las sociedades humanas y en el resultante comportamiento humano social. Los hombres se ven impulsados inicialmente a cooperar entre sí por su temor común a los animales salvajes y, después, por el temor común que sienten unos de otros. Pero luego, lentamente, llegan a darse cuenta de que «no hay especie animal que sea, sin el freno del gobierno, más incapaz de concordar en multitud durante mucho tiempo, que ésta a la que pertenece el hombre» (Mandeville 1970: 81 [1982: 23]); es decir, terminan por darse cuenta de la necesidad de «leyes» que:

> se encuentren claramente concebidas como otros tantos remedios para curar y frustrar ese instinto natural de soberanía, que le enseña al hombre a mirarlo todo como si estuviera centrado en él mismo, y lo impulsa a reclamar todo aquello sobre lo que pueda poner su mano. (Mandeville 1924: 271)

Pero dada la fuerza del amor del hombre por sí mismo, esa apreciación es insuficiente para la realización sostenida y eficaz de lo que se necesita —y lo sigue siendo aun cuando las «leyes» en cuestión se dicten y se respalden

en instituciones y prácticas basadas en el temor humano al castigo—. Los temores fundados en el amor por sí mismo no producirán formaciones sociales estables. La suficiencia ha surgido —de manera vacilante y no sistemática— mediante la realización, por parte de «políticos hábiles», de la fuerza potencial del diferente instinto del gusto por sí mismo. Mandeville expresa concisamente su opinión acerca de este proceso al comienzo de la «Investigación sobre el origen de la virtud moral» (1970: 81 y ss. [1982: 23 ss.]). Los «legisladores y otros hombres sabios que se desvelaron por la institución de la sociedad» han tratado de hacer creer a la gente que es en beneficio de cada uno velar por el interés público en lugar de por lo que parece ser su interés privado. Pero «esta tarea ha sido siempre dificultosa». El único medio viable de lograrla es ofrecer alguna «gratificación» a las acciones de espíritu público: pero no es posible «otorgar tantas gratificaciones reales que satisfagan a todas las personas por cada acción individual». De este modo, los «hombres sabios» se vieron obligados a concebir una gratificación imaginaria para «el problema de la negación de sí mismo», una gratificación que no le costara nada a nadie, pero que fuera «la más aceptable recompensa para quienes la reciben»: el elogio y la adulación. Esto es, los sabios observaron que «la adulación era el argumento más poderoso que se podía usar con las criaturas humanas». Así, al elogiar las acciones de espíritu público como nobles y racionales, y al condenar las puramente egoístas como típicas de los «brutos», los sabios pudieron persuadir a la gente de controlar, por lo menos en apariencia, sus tendencias egoístas. Pero sólo en apariencia: aun cuando las personas piensan que están motivadas por la consideración del interés público, en realidad lo que las mueve es la perspectiva de una adulación que posee la agradable utilidad de confirmarles el gusto por sí mismas. Así, los hombres sabios dividieron conceptualmente «a toda la especie en dos clases que difieren entre sí de manera radical»: la primera es la de las «personas abyectas, de pensamientos viles», que siempre buscan su placer inmediato, incapaces de negarse a sí mismas, esclavizadas por la voluptuosidad, indiferentes al bien de los otros; la segunda es la de «criaturas nobles, elevadas», liberadas de sórdidos egoísmos, animadas por el bien público y que valoran sobre todo «los progresos de la mente». Dada la distinción conceptual así introducida, junto con la plasticidad del orgullo y la cuidadosa orientación de la adulación de los hombres sabios, las personas se sintieron impulsadas constantemente a demostrar —o aparentar— que pertenecían a la segunda clase, la más noble.

De esta suerte, podemos ver que, según las palabras más conocidas de Mandeville, «las virtudes morales son la prole política que la adulación

engendra en el orgullo» (1970: 88 [1982: 28]): un tipo de proceso procreativo considerado esencial para la existencia de las afecciones sociales que hicieron posible la existencia de la sociedad humana.

En lo tocante al papel desempeñado por los «moralistas y políticos», hay una evidente ambivalencia por parte de Mandeville:

> Doy esos nombres indistintamente a todos los que, tras haber estudiado la naturaleza humana, se han esforzado por civilizar a los hombres y volverlos cada vez más tratables, *ya sea* para tranquilidad de gobernantes y magistrados, *o bien* para la felicidad temporal de la sociedad en general. (Mandeville 1732b: 41, las cursivas son mías; véase 1970: 85 [1982: 26].)

Sin embargo, de acuerdo con la psicología filosófica de Mandeville, no existe aquí problema alguno: los «políticos hábiles» pueden ser tan ignorantes de las fuerzas psicológicas que realmente los mueven como el resto de la humanidad. Si observamos «las virtudes de los grandes hombres», no encontraremos más que «polvo y telarañas» (Mandeville 1970: 187 [1982: 107]).

6.2 Pompa, vanidad e impaciencia

Para Mandeville, pues, la moralidad (en uno de los sentidos del término) constituye esencialmente un fenómeno político, no sólo social. La moralidad es un artificio humano motivado por el deseo que surge de volver a los hombres «cada vez más manejables»: esto es, motivado por el deseo de ejercer un poder institucionalizado sobre los demás hombres. Así, «lo que indujo primeramente al hombre a sofocar sus apetitos y a dominar sus inclinaciones más queridas no fue ninguna religión pagana ni otras supersticiones idólatras, sino la hábil dirección de los precavidos políticos» (1970: 87 [1982: 28]). El «instinto natural de soberanía», el «amor por el dominio y ese temperamento usurpador con que nace toda la humanidad» pueden dar origen naturalmente, dentro de las circunstancias generales de la existencia humana, a un deseo muy exacerbado de «soberanía» en relación con los otros: dentro de las circunstancias generales constituidas, digamos, por los beneficios generales de vivir en sociedad y la escasez de bienes de los que se pueda echar «mano». Cabe explicar lo anterior simplemente en términos del amor por sí mismo; sin embargo, para que el deseo así producido en los «precavidos políticos» tenga alguna oportunidad razonable de verse satisfecho, entonces, dados los hechos

del amor por sí mismo y el gusto por sí mismo, se requiere algo como el artificio de la moralidad, el cual tendrá que ser propagado.

Pero solamente «algo como», pues la moralidad no es sino uno de los artificios de que disponen en este aspecto los «hábiles políticos». (En realidad, la moralidad es en sí misma un conjunto de artificios. Una laguna en la exposición de Mandeville es la falta de una consideración directa de la cuestión de por qué los políticos y los moralistas introducen en sus sujetos una moralidad y no otra.) Mandeville insiste en que los hombres pueden llegar a sentirse orgullosos de casi todo, y ofrece un buen ejemplo de esta enigmática plasticidad del orgullo: el orgullo por las hazañas de nuestros ancestros. Resulta grato, en este contexto, el hecho de que cuando Buñuel trató de caracterizar el fenómeno del machismo, lo hizo mediante un exagerado sentido de orgullo por la propia masculinidad, pues esto presupone que un sentido de orgullo por la propia masculinidad podría *no ser exagerado*. Y quizá el caso más simpático de todos sea el siguiente: un conocido mío está realmente *orgulloso* de haber compartido el mismo elevador ¡con Marcello Mastroianni! Dada esta plasticidad del orgullo, pues, otro artificio posible para el mismo fin —que en efecto ha sido creado de este modo (atropelladamente) y que en algunas sociedades ha existido incluso al lado de un apego declarado a la moralidad— es lo que podría denominarse *la ética del honor*: ese conjunto de valores prácticos cuya aceptación se pone de manifiesto con mayor claridad en la institución del duelo, en los códigos de honor militar y las reglas de los buenos modales, pero quizá también en ciertas ideas relativas a la cacería y la caballería e incluso acaso en el fenómeno que ahora se conoce como consumo conspicuo. Esta ética del honor «tiene el mismo origen» que lo que llamaré *la moralidad de la virtud o el bien*, dado que la terminología propia de Mandeville no siempre resulta coherente en este aspecto: es decir, cada una es un artificio de «hábiles políticos» concebido para volver a los hombres «cada vez más manejables» mediante la manipulación de su promiscuo instinto de gusto por sí mismos. No obstante, pese a su similitud de origen, e incluso su coexistencia ocasional, Mandeville sostiene que la ética del honor es incompatible con la moralidad de la virtud. Quienes piensan lo contrario suponen, en efecto, «que el apego más estricto al mundo no es incongruente con la promesa del hombre de renunciar a la pompa y la vanidad de éste» (1732b: 105). Tampoco parece haber muchas dudas acerca de que la intención del argumento de Mandeville no era aquí simplemente criticar a quienes ignoran esa incompatibilidad: está claro que también pensaba que la incompatibilidad entre la ética del honor y la moralidad de la virtud constituía una crítica de esa ética. Y, sin

embargo, algunas de las mismas consideraciones que condujeron a Mandeville a sostener la tesis de la incompatibilidad figuran también en su defensa de otra tesis: ¡la de la imposibilidad que tienen los hombres para ejemplificar el aspecto positivo de la moralidad de la virtud! La virtud es «inaccesible», «algo impracticable» (Mandeville 1970: 243 [1982: 152]).

En parte, pues, Mandeville es una figura bastante conocida: el moralista preocupado por denunciar la hipocresía y el autoengaño prevalecientes en la sociedad en que vive. Aunque profesan su apego a los valores cristianos, la gente muestra en su conducta que los valores realmente vigentes en su vida son muy distintos: materialistas, hedonistas, mundanos. Pero también está presente una figura quizá menos familiar: el filósofo preocupado, como se acaba de señalar, por demostrar la tesis de la incompatibilidad. Los valores verdaderamente vigentes no son sólo diferentes de los profesados sino, de hecho, incompatibles con ellos.

> La única razón de peso que puede alegarse contra el honor moderno es que contradice directamente a la religión: la una nos obliga a soportar las injurias con paciencia, el otro nos dice que, si éstas no nos ofenden, no somos dignos de vivir. La religión nos ordena dejar la venganza en manos de Dios; el honor nos obliga a no dejar la venganza a nadie más que a nosotros mismos, aun en los casos en que la ley podría hacerlo en nuestro lugar. La religión claramente prohíbe el asesinato; el honor abiertamente lo justifica. La religión prohíbe derramar sangre, bajo cualquier pretexto; el honor obliga a batirse por cualquier nimiedad. La religión se basa en la humildad y el honor en el orgullo. El modo de reconciliar estas contradicciones, vale más dejarlo a cabezas más sabias que la mía. (Mandeville 1970: 232-233 [1982: 144])

> Ser a la par bien educado y sincero resulta poco menos que una contradicción; y, por lo tanto, cuando un hombre aumenta sus conocimientos y perfecciona sus modales, debemos esperar verle al mismo tiempo ampliar sus deseos, refinar sus apetitos y desarrollar sus vicios. (*Ibid.*: 202 [1982: 119])

Que esta ética del honor es «diametralmente opuesta» (Mandeville 1732b: 77) a la moralidad de la virtud se confirma además por otras varias consideraciones. Una de ellas se ilustra con el hecho de que los buenos modales, por ejemplo, son compatibles con la insinceridad, mientras que la virtud no lo es: en general, en la ética del honor, a la cuestión de la motivación —*la* cuestión para la moralidad de la virtud— se le concede, a lo sumo, un puesto secundario en relación con el del desempeño público. En contraposición con la ética del honor, desde el punto de vista de la moralidad de la virtud «es imposible juzgar las acciones de los hombres, a

menos que estemos muy bien enterados del principio y del motivo que los hacen obrar» (Mandeville 1970: 91 [1982: 31]). Y esto se relaciona, desde luego, con el hecho de que el honor sea más prontamente detectado por los demás que la virtud:

> Un hombre puede ser justo y casto y, sin embargo, ser incapaz de convencer al mundo de que lo es; pero puede provocar una pelea y mostrar que se atreve a luchar cuando le place, en especial si tiene trato con hombres de armas. Cuando el principio del honor se tenía en alta estima, la vanidad y la impaciencia deben haber impulsado siempre a los más orgullosos y atrevidos a buscar oportunidades de señalarse, con el fin de que se los considerase hombres de honor. Esto ocasionaba naturalmente disputas y luchas. (Mandeville 1732b: 63)

En realidad, a los hombres de honor las «oportunidades de señalarse» les resultan esenciales para el *objetivo* de su práctica:

> Un hombre virtuoso no espera ningún reconocimiento de los demás; y si éstos no creen que sea virtuoso, su tarea no es obligarlos a creerlo. Pero un hombre de honor posee la libertad de proclamarse abiertamente como tal, y pedirle cuentas a todo aquel que se atreva a dudarlo. (*Ibid.*: 44)

> La recompensa de la gloria [...] consiste en una felicidad superlativa, que el hombre consciente de haber realizado una acción noble se goza en el amor propio, mientras piensa en los aplausos que espera. (Mandeville 1970: 90–91 [1982: 30])

En contraposición, pues, con la moralidad de la virtud, la ética del honor crea para sus practicantes «oportunidades de señalarse» y autoriza luego la exigencia de «reconocimiento» de los individuos así señalados. Y, por último, hay una incompatibilidad práctica entre la ética del honor y la moralidad de la virtud: «Las buenas maneras no tienen nada que ver con la virtud o la religión; en lugar de extinguirlas, más bien inflaman las pasiones» (*ibid.*: 112 [1982: 48]).

6.3 ¿Un caballo que no baila?

Desde cierto punto de vista, la ética del honor es «un logro mayor» que la moralidad:

> Porque la una se encuentra adaptada más hábilmente a nuestra constitución interna. A los hombres les rinde mejores frutos su apego al honor que su

> apego a la virtud: el primero exige una menor negación de sí mismo, y las recompensas que se reciben por esa pequeñez no son imaginarias, sino reales y palpables [...] La invención del honor ha sido muchísimo más benéfica para la sociedad civil que la de la virtud, y ha respondido mucho mejor al fin para el cual fueron inventados. Pues desde que la noción del honor fue aceptada entre los cristianos, siempre ha existido, en la misma cantidad de personas, veinte hombres de verdadero honor por uno de verdadera virtud. (Mandeville 1732b: 43-44)

Ese juicio sobre los artificios en cuestión está explícitamente relativizado al «fin para el cual fueron inventados», y resulta por lo tanto compatible con el otro juicio de Mandeville en el cual la moralidad de la virtud se usa para criticar la ética del honor. Tampoco debemos permitir que nos descarríe la observación de Mandeville de que: «Lo único de peso que puede decirse contra el honor moderno es que se opone directamente a la religión.» Tal vez sea lo único que pueda decirse, pero al leer lo que Mandeville escribe sobre el duelo, tenemos que admitir que, para él, el peso de esa única cosa podría ser considerable.

No obstante, debe surgir la sospecha de que, para Mandeville, las probabilidades de que nos topemos con un hombre «de verdadera virtud» en lugar de con uno «de verdadero honor» van probablemente mucho más allá de veinte contra uno —no es sólo que los hombres virtuosos «hasta ahora no hayan abundado en el mundo» (1970: 92 [1982: 31])—. A diferencia del «verdadero honor», la «verdadera virtud» exige una completa «renunciación» (según «la generalidad de los moralistas y filósofos» (*ibid.*: 329)). Aun así, para Mandeville todas las motivaciones naturales de los hombres emanan del «amor por sí mismo» o del «gusto por sí mismo», o de alguna combinación de ambos; así pues, dejando a un lado la supuesta posibilidad de la gracia divina, se sigue que la «verdadera virtud» es para los hombres una imposibilidad. De aquí la objeción de Law a Mandeville:

> Primero. Considera usted al hombre *meramente* como un *animal* que, como los otros animales, nada tiene que hacer salvo seguir sus apetitos.
>
> Segundo. Considera usted que al hombre lo arrancan de su estado natural los engaños y adulaciones pergeñados por los moralistas, y pretende estar muy seguro de que *«las virtudes morales son los retoños políticos que la adulación procreó del orgullo»*.
>
> De suerte que aquí el hombre y la moralidad se destruyen conjuntamente. Se declara que el hombre es únicamente un *animal*, y la moralidad, una impostura. Según esta doctrina, calificar a un hombre de deshonesto es volverlo exactamente tan criminal como un caballo que no baila. (Law 1845: 3-4)

Desde luego, hay aquí algunos malentendidos terribles, si no tergiversaciones deliberadas. Pero intentemos ver si se presenta en este punto algún reparo que valga la pena, pertenezca o no a Law.

En contra de quienes, como Law, desearían contemplar que el hombre ha sido creado a imagen y semejanza de Dios, Mandeville, el «amante de la experiencia» (1711: xi), insistía en la necesidad de contemplarlo en su «desnudez», por desagradable que pudiera resultar soportar esto. Ambrose Bierce definió una vez al cinismo como un defecto de la visión que nos obliga a ver el mundo tal como es, en lugar de como debiera ser. Con este «cinismo» acerca de la naturaleza del hombre —«Una de las principales razones del porqué son tan raras las personas que se conocen a sí mismas, radica en que la mayoría de los escritores se ocupa de enseñar a los hombres cómo deberían ser, sin preocuparse casi nunca por decirles cómo son en realidad» (Mandeville 1970: 77 [1982: 22])—, Mandeville se vio impulsado a destacar el predominio en los asuntos humanos de la hipocresía y el autoengaño, la fuerza del deseo obsesivo de orgullo y estima que alientan los hombres, y la sorprendente disparidad general entre lo que los hombres dicen y lo que hacen: en suma, se vio llevado a reconocer «la bajeza y la deformidad reales» (1732a: 48) de la naturaleza humana.

> [L]o que *Emilia* piensa de sí misma es digno de atención. Está segura de que su prostitución no emanó jamás de la concupiscencia, sino de la necesidad, *ergo* no hay pecado. El daño que hace con su maledicencia lo atribuye a la aversión que siente por el vicio. Cuando reflexiona en las horas que pasa en la iglesia y leyendo, y piensa entonces en el testamento que ha hecho, se elogia a sí misma por haber cumplido con todos los deberes cristianos, y su conciencia está enteramente limpia. ¿No es extraño que *Emilia*, con todo su ingenio, no haya sospechado nunca que es una mujer enferma, y que hasta el momento no se percate de que la envidia y la vanidad son sus vicios predilectos? (Mandeville 1720: 35)

> Algunos embaucan al mundo, y quisieran que se les crea que creen lo que realmente no creen; pero la mayoría se engañan a sí mismos, sin tener en cuenta ni comprender plenamente en qué consiste el creer. (Mandeville 1970: 186 [1982: 107])

> Me declaro dispuesto a rendir homenaje a la virtud, dondequiera que pueda encontrarla, con la condición de que no se me obligue a admitir como tal ninguna donde yo no pueda encontrar abnegación, ni juzgar los sentimientos de los hombres por sus palabras cuando tengo delante de mis ojos los actos de su vida. (*Ibid.*: 174 [1982: 97])

Sin duda, Law deseaba objetar tanto el carácter empírico y naturalista de esta conjetura, como su contenido; pero en ambos aspectos me parece que Mandeville es irrebatible. Luego, sin embargo, sobre esta convincente base empírica, Mandeville intenta construir una teoría excesivamente simplista de la motivación humana. Al proceder de este modo, cae en las trampas habituales: desde la de concebir a la mente como si tuviera cierta similitud con el cuerpo en sus operaciones, pasando por la de soslayar la variedad de tipos del deseo humano, hasta la de no tener en cuenta la variedad de los objetos del deseo humano. Dada la teoría de la motivación así construida, junto con sus ideas acerca de lo que exige la moralidad de la virtud, Mandeville se ve entonces llevado a negar la posibilidad efectiva de una «verdadera virtud» para la raza humana natural. Otras objeciones razonables se suscitan con respecto al método que sigue Mandeville para tratar de sostener esa negación. En primer lugar, sus razones en apoyo de la negación descansan en los no bien explicados «términos técnicos» que despliega en su teoría de la naturaleza humana, el «amor por sí mismo» y el «gusto por sí mismo». Y en segundo lugar, su negación también descansa en una afirmación no defendida acerca de la primacía, para la evaluación moral de las acciones, de una teoría apenas articulada sobre lo que sería la motivación moral:

> Pero no hay que juzgar a los hombres por las consecuencias que puedan seguir a sus acciones, sino por los hechos mismos y por los motivos que parecen ser causas de ellas. (Mandeville 1970: 119 [1982: 53])
>
> No veo [...] abnegación alguna sin la cual no puede existir la virtud. (*Ibid.*: 177 [1982: 99])
>
> Las pasiones pueden hacer bien por casualidad, pero sólo en vencerlas consiste el mérito. (*Ibid.*: 107 [1982: 44])

Ésas no son objeciones de Law. Aparte de las que implica su acusación de que Mandeville contempla al hombre «*meramente* como un *animal*», su principal objeción parece ser que la conclusión ulterior que Mandeville quiere que saquemos es que deberíamos rechazar la moralidad de la virtud. Supuestamente, la moralidad ha sido «destruida»: y ello no sólo en el sentido de que en su parte positiva la «institución» esté vacía en relación con las acciones humanas, sino también en el otro sentido de que se nos invita, de manera consciente, a abandonar sus mismos términos de referencia.

Pero la conclusión ulterior que Mandeville oficialmente desea que saquemos es muy diferente:

> Puesto que es imposible servir a Dios y a Mamón, no tardaré en escoger ningún placer temporal amerita correr el riesgo de ser desdichado por toda la eternidad, y déjese a quien quiera trabajar para engrandecer a la nación, yo me ocuparé de fines más elevados y velaré por mi propia alma. (Mandeville 1732a: 22)

Ahora bien, si se le preguntara qué piensa de la «verdadera virtud», apenas bastaría que Mandeville ofreciera la respuesta de Gandhi a la misma pregunta acerca de la «civilización británica»: «Ésa sería una buena idea.» En este pasaje, Mandeville está señalando un motivo no moral para el continuo despliegue de los términos de referencia de la moralidad de la virtud: no moral porque, primero, de acuerdo con su propia teoría, ciertamente vaga, sobre la motivación moral, la perspectiva de que existan recompensas y castigos divinos en interacción con un hedonismo racional no podría producir acciones de «verdadera virtud»; y segundo, a la luz de su idea de la imposibilidad de que los seres humanos reales tengan motivaciones morales, toda motivación efectiva de ese continuo despliegue tiene que ser no moral. Pero surge la cuestión de cuál es exactamente la *manera de desplegar* los términos de referencia de la moralidad de la virtud que a Mandeville le interesa aprobar. Pueden distinguirse aquí por lo menos tres posibilidades. En primer lugar, aun cuando la «verdadera virtud» sea algo imposible para los seres humanos reales, no se sigue que todas las acciones humanas sean viciosas ni que todas las acciones viciosas lo sean en igual grado. El apego a la moralidad de la virtud no puede desembocar, *ex hypothesi*, en acciones de «verdadera virtud»: pero, pese a todo ello, dicho apego podría desempeñar, no obstante, cierto papel en la selección de los medios moralmente neutrales o menos malos para llegar a un determinado fin ciertamente vicioso, e incluso en la selección de un fin moralmente neutral o menos malo en vez de otro más vicioso. En segundo lugar, aun cuando la «verdadera virtud» sea imposible para los seres humanos reales, no se sigue que sea incoherente el proyecto de tratar de alejarse uno mismo del vicio, de tratar, por ejemplo, de fortalecer las motivaciones menos malas o incluso moralmente neutrales. Y en tercer lugar, la imposibilidad de una «verdadera virtud» no establece la incoherencia del proyecto de intentar alcanzar la coincidencia externa entre las propias acciones y las que serían ejemplo de la «verdadera virtud».

Una cosa está clara: la afirmación de Mandeville de que unos «hábiles políticos» crearon la moralidad de la virtud no debe confundirse con la idea de que esos políticos se las ingeniaron para volver a los hombres realmente virtuosos. Lo que los políticos crearon fue *la imagen de un ideal*

imposible, una imagen cuya existencia podía redundar en su propio beneficio (y en el de otros) por medio de la amplificación consecuente de las clases de acciones egoístas, vanagloriosas, que los hombres tenían a su alcance. Al adular a sus súbditos, al menos algunas veces, como ejemplos de ese ideal imposible, estos «hábiles políticos» pueden hacerlos «cada vez más manejables»: pues el orgullo placentero que potencialmente sienten los súbditos acerca de su supuesta virtud moral constituye una fuerza lo bastante poderosa para preponderar sobre sus otros instintos antisociales. Y si los mismos «hábiles políticos» hubieran llegado a abrigar dudas acerca de la eficacia general de la moralidad de la virtud en este respecto, habrían podido recurrir a otras invenciones, como la ética del honor, calculadas para reforzar el mismo efecto general deseado de volver a los hombres «cada vez más manejables»; si así sucede, dados los presuntos fines de los «hábiles políticos», no es preciso que les preocupe ningún asunto relativo a la inconsistencia entre la moralidad de la virtud y las demás invenciones empleadas. No obstante, surge un peligro doble para esos políticos: el primero es que, como resultado de los esfuerzos de alguien como Mandeville, la incongruencia entre las diversas invenciones se haga del conocimiento público; y el segundo, que, como resultado de los esfuerzos de alguien como Mandeville, la imposibilidad de la «verdadera virtud» se haga del conocimiento público. Lo primero puede sacudir al sistema, lo segundo puede «destruirlo». Pues si la «verdadera virtud» es imposible, la adulación de los sujetos en cuanto a su ejemplificación carece siempre de fundamento y el orgullo sentido en cuanto a esa ejemplificación jamás está justificado; y si esto llegara a saberse públicamente, es difícil entender cómo la invención de la moralidad de la virtud podría continuar funcionando para los fines de los «hábiles políticos». De suerte que, aunque los esfuerzos de Mandeville no amenacen «destruir» la moralidad de la virtud en sí, sí amenazan destruir el *uso* que de ella hacen los «hábiles políticos».

Supongamos que, en efecto, así ha ocurrido: ¿cuáles son las reacciones posibles de los «hábiles políticos»? Un conjunto de reacciones, quizá sólo accesible a los más hábiles, está constituido por las tentativas por trabajar con la dimensión negativa de la moralidad de la virtud, la dimensión que la discusión de Mandeville ha dejado en realidad intacta. De modo que una reacción consistiría en continuar con la adulación de los súbditos, pero ahora señalando que son menos viciosos que los demás; otra sería la de continuar con esa misma adulación, pero ahora aduciendo el «verdadero mérito moral» de los súbditos, que se pone de manifiesto en su

capacidad de sentir remordimiento y culpa por sus propios vicios (de manera que esos sentimientos se conviertan, de modo bastante paradójico, en nuevos objetos del plástico orgullo). Pero hay otro conjunto de reacciones que tiene mayores probabilidades de producirse entre los políticos menos hábiles: el constituido por las tentativas de sustituir la moralidad de la virtud por otros grupos de presuntos valores que no se enfrentan al problema de la imposibilidad. Los candidatos nos resultan ahora demasiado familiares: el racismo, el nacionalismo, el colonialismo interno, el sexismo y hasta la idea misma de «moralidad sexual». Dada la plasticidad del orgullo, cada uno de éstos permite al político adular por lo menos a algunos de sus súbditos y puede de este modo suscitar, en los súbditos en cuestión, sentimientos de orgullo por su superioridad, sentimientos que inducen a la aquiescencia. Y no es necesario que ninguna de estas invenciones le cueste algo al político —siempre que éste pertenezca a la raza correcta, a la nacionalidad, la clase social y la casta correctas, al sexo y las preferencias sexuales correctas—. (Y si él no es así: *¡cuidado!*)

6.4 ¿Todas las abejas han perdido su aguijón?

Quizá no hemos considerado todavía la razón que se aduce con más frecuencia para apoyar la acusación de «inmoralismo» contra Mandeville. Se trata de la afirmación, que es supuestamente la moraleja de la fábula de las abejas, de que la existencia de los beneficios públicos, como la riqueza y la grandeza nacionales, dependen del carácter vicioso del comportamiento personal —es decir, individual—, de suerte que si los hombres llegaran, *par impossible*, a llevar vidas de «verdadera virtud», esos beneficios públicos de la vida social desaparecerían. Por consiguiente, según los críticos de Mandeville, éste *recomendaba* que el comportamiento privado fuera vicioso.

No se considerarán aquí los posibles argumentos en favor de la tesis de que los vicios privados redundan en beneficios públicos; y no perderé el tiempo con los intentos más recientes de extender la aplicación del modelo del «hombre económico» que prevalece entre los economistas contemporáneos del *laisser-faire* para «explicar» áreas muy distintas de la conducta humana. (De hecho, Mandeville percibió aquí una parte del error: «Las matemáticas llegan a ser el único estudio valioso y se usan en todo, aun donde es ridículo», 1970: 323 [1982: 214].) Aquí me limitaré a indicar algunas de las motivaciones que sospecho que condujeron a Mandeville a idear su fábula y que, si realmente fueran suyas, estarían cerca de eximirlo del cargo de «inmoralismo».

Creo que está claro que la motivación de Mandeville para idear su fábula fue, en parte, responder a una objeción obvia a su explicación de la naturaleza humana: si el hombre fuera una criatura tan viciosamente egoísta y antisocial como esa explicación lo supone, ¿cómo podría llegar a existir el fenómeno de la sociedad humana? El hecho de que la réplica a esa objeción constituya una parte de la motivación de Mandeville lo sugiere la portada de la segunda de las dos ediciones de 1714: «La fábula de las abejas, o vicios privados, beneficios públicos. CONTIENE varios discursos para demostrar que las flaquezas humanas, *durante la degeneración de la* HUMANIDAD, pueden redundar en provecho de la SOCIEDAD CIVIL y ser empleadas para ocupar el lugar de las *virtudes morales*» (citado en Monro 1975: 189, las cursivas y las versalitas son del original). De suerte que, pese a la naturaleza esencialmente viciosa y antisocial del hombre, el hecho de que exista la sociedad humana puede, en efecto, explicarse. Mandeville no estaba obligado a negar ese hecho, como tampoco sus críticos deberían haber negado esa naturaleza.

El mismo interés por ser «cínico» en el sentido de Bierce se relaciona con otra parte de la motivación que llevó a Mandeville a idear su fábula. Si encaramos alguna vez los hechos, Mandeville piensa que nos veremos llevados a reconocer el desorden moral del mundo en que vivimos: pues en este mundo lo bueno y lo malo se encuentran tan mezclados entre sí como los ingredientes de un «bol de ponche».

> El vulgo miope, en la cadena de las causas no suele ver más allá del eslabón inmediato; pero los que pueden ensanchar su visión y entregarse al placer de echar una mirada a los acontecimientos concatenados, podrán ver en cien lugares cómo el *bien* emerge y pulula del *mal* con tanta naturalidad como los polluelos de los huevos. (Mandeville 1970: 123 [1982: 56])

> En la moralidad, lo mismo que en la naturaleza, nada existe en las criaturas tan perfectamente bueno que no pueda resultar perjudicial para nadie de la sociedad, ni tan totalmente malo que no pueda ser benéfico para una parte u otra de la Creación [...] (*Ibid.*: 369 [1982: 247])

> En reponer lo que se pierde y destruye por el fuego, las borrascas, los combates navales, los sitios y las batallas, consiste una gran parte del movimiento mercantil [...] (*Ibid.*: 361 [1982: 241])

Los ejemplos pueden multiplicarse (y Mandeville los multiplica). Pero si alguna vez reconocemos este carácter maniqueo del mundo en que

vivimos, nos veremos llevados a aceptar entonces la necesidad de las decisiones —como, por ejemplo, la de si vale la pena pagar cierto precio o la de si ha de renunciarse a un beneficio potencial—. El rechazo a hacer frente al desorden moral de este mundo y el consecuente rechazo a hacer frente a la necesidad de tomar decisiones constituyen, por lo menos, uno de los blancos principales de la fábula de Mandeville.

Quizá en uno de sus momentos menos perspicaces, Hume preguntó lo siguiente: «¿No es muy incongruente que un autor afirme en una página que las distinciones morales son invenciones de los políticos en bien del interés público, y que en la siguiente página sostenga que el vicio es ventajoso para el público?» (Hume 1963a: 163). Aquí la confusión resulta clara: lo ventajoso es la existencia de una *imagen* de un ideal imposible, no la existencia de un comportamiento que ejemplifique la «verdadera Virtud». Pero tal vez la objeción real de Hume era otra: «En cualquier sistema moral, parece poco menos que una contradicción en los términos hablar de un vicio que, en general, resulta benéfico para la sociedad» (*ibid.*).

Ahora bien, acabamos de ver que uno de las preocupaciones de Mandeville era enfatizar la necesidad que hay en este caso de tomar decisiones; también hemos visto (pp. 219–220) que su propia decisión, por lo menos desde un punto de vista teórico, era clara: «Aunque haya indicado el camino de la grandeza mundana, siempre he preferido, sin la menor vacilación, el que conduce a la virtud» (Mandeville 1970: 241 [1982: 151]). La cuestión es ahora que esa declaración de su decisión debería entenderse como una llamada de atención hacia la posibilidad de que la consideración de las relaciones entre el comportamiento privado y los «beneficios públicos» podría suscitar dudas razonables acerca de si tales «beneficios públicos» son realmente benéficos, si realmente valen la pena en el mundo tal como es. La determinación de Mandeville de mostrarse «cínico» con respecto al desorden evaluativo de este mundo lo condujo a apreciar la complejidad de lo que aquí se ha denominado «deliberación protopráctica» (p. 32): una complejidad tan desdeñada por sus contemporáneos optimistas morales como por muchos de nuestros teóricos de la decisión contemporáneos.

6.5 El amplio manto de la sociabilidad

Para terminar esta discusión sobre nuestro «misionero del reino de las tinieblas», quiero plantear una cuestión general: si la explicación que ofrece Mandeville sobre los orígenes del pensamiento moral fuera más

o menos correcta, ¿cómo podríamos esperar que el pensamiento y el discurso morales funcionaran realmente? Ya hemos considerado el funcionamiento probable de ese discurso en boca de «hábiles políticos», tanto antes como después del reconocimiento general de los resultados de los esfuerzos de alguien como Mandeville. ¿Qué pasa ahora con el hombre ordinario dentro de la sociedad? Tras regodearse al principio con la adulación que hacen los políticos de su virtud moral, querrá todavía aumentar la cantidad de adulación: ningún elogio es suficientemente bueno. («Pocos hombres pueden ser persuadidos de que ganan demasiado por aquellos a quienes venden algo, por extraordinarias que sus ganancias sean» (Mandeville 1970: 113 [1982: 49]).) La manera más sencilla de aproximarse al límite imposible sería mediante la *auto*adulación. No obstante, por razones fácilmente comprensibles dentro de los términos de la propia explicación de Mandeville, la *modestia* (en una de sus «acepciones») será considerada una de las virtudes sociales cardinales (*ibid.*: 101 y ss. [1982: 40 y ss]). Nuestro hombre en la sociedad puede tratar entonces de «simular» esa virtud, de «encubrir con destreza» (*ibid.*: 156 [1982: 82]) su orgullo, y diseña esta simulación estrictamente para engañar a los demás, de suerte que éstos se avengan a la persecución de su meta: así que puede esperar recibir, gracias a esto, una adulación directa de los demás por su virtuosa «modestia», o bien provocar con esta actitud la adulación de los otros en términos de la «verdadera virtud» que se esconde bajo esa «modestia». Y si no logra entonces la adulación deseada de esos otros, le queda todavía un recurso adicional: desdeñar la aprobación de esos individuos como si ésta no mereciera la pena, para luego buscar a otros que compartan, o que por lo menos parezcan compartir, la alta opinión que tiene de sí mismo. Si le resulta demasiado difícil encontrar a esos otros —o quizá incluso si le resulta *demasiado fácil* encontrarlos—, es probable que nuestro hombre, dentro de las convenciones de esta «modesta sociedad», recurra entonces a otro mecanismo: tratar de obtener una adulación *implícita* a través de la crítica de los defectos morales *de otros*, una crítica concebida para realzar la conciencia, sea la propia o la de los demás, de su propia «virtud» comparativa («con la misma injusticia con que pensamos tan favorablemente de nosotros mismos, solemos pensar mal de nuestros vecinos» (Mandeville 1970: 158 [1982: 84])). Pero, dados los instintos dominantes del «amor por sí mismo» y el «gusto por sí mismo», sería sumamente improbable que ocurriera cualquier tentativa seria de autocrítica moral y cualquier tentativa seria de perfeccionamiento moral de la propia persona: e incluso si *pareciera* que tales fenómenos ocurren, probablemente no serían más que otras estratagemas concebi-

das para suscitar el pensamiento adulador, en los demás o en nuestro hombre mismo, de que está demostrando así su propio valor moral, su propia sensibilidad y capacidad para la autocrítica moral. Todo esto exige, desde luego, un considerable autoengaño.

Ahora bien, si los resultados de la obra de alguien como Mandeville llegan a ser conocidos en un nivel general, las posibilidades efectivas de autoengaño se verán, por lo menos, muy reducidas: generalmente, la simple hipocresía tomará su lugar.

> Avergonzados de las muchas flaquezas interiores que padecen, todos los hombres se esfuerzan por ocultarse unos a otros la fealdad de su desnudez, envolviendo el verdadero sentir de sus corazones en el amplio manto de la sociabilidad, con la esperanza de esconder sus sucios apetitos y la deformidad de sus deseos bajo la apariencia de un sincero interés por el bien público; mientras que en su fuero interior tienen conciencia de su invencible afición a sus concupiscencias favoritas y de su absoluta incapacidad de transitar por el arduo y fragoso sendero de la virtud. (Mandeville 1970: 244 [1982: 153])

Tal hombre en sociedad no puede enorgullecerse de su «verdadera Virtud». Pero además del rechazo de los mismos términos de referencia de la Moralidad de la Virtud, le queda todavía otro curso de acción: el de enorgullecerse de ser mejor —es decir, menos vicioso— que Los Demás. Ahora la censura moral de los demás se vuelve esencial para su propia adulación comparativa. Los juicios negativos sobre la moralidad de los otros será la norma: serán el único medio de que dispone para definirse a sí mismo en términos moralmente aceptables, la condición misma para hacerlo así. Y, pese a conocer los resultados de los esfuerzos de alguien como Mandeville, su «amor por sí mismo» y su «gusto por sí mismo» instintivos continuarán manteniendo en grado mínimo la probabilidad de una autocrítica moral seria y de un intento serio de autoperfeccionamiento moral. En realidad, si parece que tales fenómenos ocurren, probablemente no serían sino una manifestación más del «gusto por sí mismo» —el que está implicado en su idea de que él, *a diferencia de otros*, es por lo menos capaz de reconocer sus propios defectos morales y de tratar de hacer algo al respecto—.

Alasdair MacIntyre sostiene que es el hecho de que «la moralidad se encuentre hoy en un estado de grave desorden» lo que explica por qué el significado del discurso moral puede ocultar su uso (MacIntyre 1981: 66). De los escritos de Mandeville he extraído una descripción alternativa de ese uso, junto con una posible explicación alternativa de la manera

como ese uso se ha producido y, por lo tanto, de la manera como el significado del discurso moral pudo ocultar ese uso. Estas alternativas no revelan nada, me parece, acerca del estado de la moralidad, acerca de *su* orden o desorden; sin embargo, la descripción alternativa sí nos aproxima a ciertas ideas sobre el funcionamiento efectivo del pensamiento y el discurso morales que se encuentran en los escritos de un pensador mejor conocido, a quien paso ahora a considerar.

6.6 Una hoja de parra moral

Para Nietzsche los peores lectores «son aquellos que proceden como soldados durante el saqueo: toman unas cuantas cosas que pueden usar, mancillan y confunden el resto, y lanzan blasfemias contra todo» (Nietzsche 1966b: 155). Aquí voy a correr conscientemente el riesgo de hallarme por lo menos a una tercera parte del camino de convertirme en semejante lector: dados mis muy particulares propósitos en este texto, tomaré unas cuantas cosas de los escritos de Nietzsche que creo que puedo usar. Pero podría haber, sin embargo, un efecto incidental en cuanto a un *ordenamiento* limitado de algunos de los pensamientos nietzscheanos sobre la moralidad.

Cualquier lector de los escritos de Nietzsche se enfrentará a cierta dificultad que les confiere un interés especial, dados mis propósitos actuales. La dificultad estriba en saber, en diversos pasajes, cuál es exactamente su objetivo. En ocasiones resulta claro que el pretendido objeto de sus críticas es un código moral muy específico: con mucha frecuencia, ciertas creencias cristianas características o incluso «toda la moralidad cristiana europea», pero también, a veces, varias creencias socialistas, y aun de vez en cuando algunos elementos del código moral que predominaba exclusivamente en la cultura alemana de su época. En otros momentos resulta evidente que el pretendido objeto de las críticas nietzscheanas es, por así decirlo, la clase sacerdotal: quienes pretenden ser autoridades de la verdadera moralidad y promotores de ella, e incluso esos moralistas jactanciosos que renuncian al mundo por las vías del ascetismo. Nuevamente, a veces resulta claro que el supuesto objeto de sus críticas es una teoría filosófica de la moralidad específica: con más frecuencia, las teorías kantianas, pero también en ocasiones, por ejemplo, las teorías utilitaristas (en especial, las planteadas por los ingleses). Pero a veces es obvio que Nietzsche tenía en mente un objetivo de mayor envergadura para sus críticas —la «moralidad misma»—. Así, de «nuestros filósofos morales» escribe lo siguiente:

> [N]o llegaron a ver en absoluto los auténticos problemas de la moral — los cuales no emergen más que cuando se realiza una comparación de *muchas* morales. Aunque esto suene muy extraño, en toda «ciencia de la moral» ha venido *faltando* el problema mismo de la moral: ha faltado suspicacia para percibir que aquí hay algo problemático. (Nietzsche 1966a: 287–288 [1980a: 114])

Todo ello parece, en efecto, bastante claro; pero hay otras dos consideraciones que presentan al lector de Nietzsche una dificultad importante. En primer lugar, en diversos pasajes no queda nada claro cuál supone Nietzsche que es exactamente el objetivo de sus críticas; y en segundo lugar, aun cuando eso se aclare, con frecuencia no queda empero nada claro que Nietzsche identifique correctamente el blanco alcanzado en realidad por esas críticas. A la luz de esa dificultad y dentro del contexto conformado por los particularísimos propósitos de este ensayo, me propongo explorar la siguiente estrategia: considerar varias críticas de Nietzsche *como si* tuvieran por objetivo la «moralidad misma», con la esperanza de que la evaluación de esas críticas entendidas de esa manera contribuya al cometido de comprender qué es la «moralidad misma» o, por lo menos, al de comprender *lo que no es*.

«No existen», nos dice Nietzsche, «fenómenos morales, sino sólo una interpretación moral de fenómenos [...]» (1966a: 275 [1980a: 99]). Esa afirmación se explica quizá un poco en el siguiente pasaje:

> [*N*]*o existen hechos morales*. El juicio moral tiene en común con el religioso el creer en realidades que no lo son. La moral es únicamente una interpretación de ciertos fenómenos, dicho de manera más precisa, una interpretación *equivocada*. El juicio moral, lo mismo que el religioso, corresponde a un nivel de ignorancia en el que todavía falta el concepto de lo real, la distinción entre lo real y lo imaginario: de tal manera que, en ese nivel, la palabra «verdad» designa simplemente cosas que hoy nosotros llamamos «imaginaciones». (Nietzsche 1968b: 55 [1980b: 71])

Un poco más adelante Nietzsche dice, con respecto a la moralidad, que «uno ya debe saber *de qué* trata para sacar provecho de ella», y saber esto exige, para Nietzsche, comprender que, como la religión, la moralidad cae enteramente

> bajo la *psicología del error*: en cada caso particular son confundidos la causa y el efecto; o la verdad es confundida con el efecto de lo *creído* como verdadero; o un estado de conciencia es confundido con la causalidad de ese estado. (*Ibid.*: 52–53 [1980b: 68])

La teoría nietzscheana del error es una manifestación notablemente rica de profundas intuiciones psicológicas. Aquí me limitaré a describir en términos muy esquemáticos la aplicación que él hace de esa teoría al «problema de la moralidad». El hombre, dice Nietzsche, «prefiere querer *la nada* a *no querer*» (1956b: 299 [1981: 186]); esa propensión interactúa con un rasgo característico de ciertas «personas pequeñas» para producir una proyección errónea sobre el mundo: «Quien no sabe introducir su voluntad en las cosas, introduce en ellas al menos un *sentido*: es decir, cree que hay ya allí dentro una voluntad (principio de la "fe")» (1968b: 24 [1980b: 32]). Esa proyección equivocada, y sus constantes acompañantes que son el error y el autoengaño en la forma en que los agentes entienden sus propias acciones, pueden ser objeto de muchos otros desarrollos, pueden desembocar en «*muchas* moralidades». Así, Nietzsche considera, por ejemplo, que la moral cristiana está característicamente arraigada en la debilidad, el temor y en ese peculiar tipo de maldad que él denomina *resentimiento*. Philippa Foot expresa del siguiente modo el punto de vista de Nietzsche:

> Quienes cultivan la humildad y las otras virtudes propiciatorias para encubrir su debilidad alimentan un resentimiento envidioso en contra de los que son más fuertes que ellos. Quieren venganza por su inferioridad y tienen un profundo deseo de humillar y hacer daño. El deseo de castigar le parece a Nietzsche uno de los signos más evidentes de esta malicia escondida, y ve la idea del libre albedrío y de la responsabilidad como algo inventado por aquellos que deseaban infligir un castigo. (Foot 1973: 83 [1994: 103])

Por lo tanto, Nietzsche comenta de la doctrina de «no buscar el *propio provecho*» que «no es más que la hoja de parra moral para tapar un hecho completamente distinto, a saber, fisiológico: "yo ya no sé *encontrar* mi provecho"» (1968b: 87 [1980b: 108]). La utilidad de esa hoja de parra depende de la capacidad de autoengaño que tenga el agente, la capacidad de ocultarse sus errores motivados. Pero el costo de sustentar una moralidad cristiana no se restringe, ni siquiera para las «personas pequeñas», a esos errores motivados y autoengaños:

> [El hombre] se tensa en la contradicción «Dios y demonio», y todo no que se dice a sí mismo, a la naturaleza, a la naturalidad, a la realidad de su ser, lo proyecta fuera de sí como un sí, como algo existente, corpóreo, real, como Dios, como santidad de Dios, como Dios juez, como Dios verdugo, como más allá, como eternidad, como tormento sin fin, como infierno, como inconmensurabilidad de pena y culpa. (Nietzsche 1956b: 226 [1981: 106])

Sin embargo, el primer motor de la generación de la moralidad cristiana se encuentra también en la generación de todas las otras moralidades: «Ésa es vuestra voluntad, sapientísimos, una voluntad de poder: y ello aunque habléis del bien y del mal y de las valoraciones» (Nietzsche 1961: 136 [1979a: 169]). Todas las moralidades, con su distintivo contraste entre el bien y el mal (*böse*), son productos de la voluntad de poder que interactúa con diversos factores deformantes: pues los productos, las moralidades, sirven precisamente para *enmascarar* esa voluntad de poder. De esto «trata» la moralidad. Y a quienes son capaces de liberarse de la influencia de los diversos factores deformantes no les queda sino una esperanza: sólo si se unen conscientemente a esta voluntad de poder, sólo si aceptan conscientemente el egoísmo que «forma parte» «de la esencia del alma aristocrática» (Nietzsche 1966a: 405 [1980a: 234]), sólo si se experimentan a sí mismos como «determinadores de los valores» (*ibid.*: 395 [1980a: 223–224]), sólo si se reconocen como «creadores de valores» (*ibid.*: 395 [1980a: 224]), hay alguna posibilidad de que incluso tales hombres escapen de los errores inherentes a la moralidad. Únicamente de esta manera tendrán alguna posibilidad de convertirse, como Zaratustra, en «un preludio de actores mejores» (Nietzsche 1961: 226 [1979a: 289]).

6.7 La congruencia inglesa

La explicación del supuesto error es notable; pero, ¿por qué se supone que aquí hay un error que debe explicarse? ¿Por qué, por ejemplo, Nietzsche niega la existencia de hechos morales, de fenómenos morales? Comienza a surgir una posibilidad en las siguientes observaciones expresadas al discutir a George Eliot:

> Se han desprendido del Dios cristiano, y creen tanto más tener que conservar la moral cristiana: ésa es congruencia *inglesa* [...] Cuando uno abandona la fe cristiana pierde el *derecho* a apoyarse en la moral cristiana. Ésta *no* es en modo alguno algo evidente de por sí [...] (Nietzsche 1968b: 69 [1980b: 88])

Esto parece una impugnación dirigida a una familia particular de moralidades. Sin embargo, un poco más adelante, Nietzsche dice lo siguiente:

> Si de hecho los ingleses creen que ellos saben de por sí, «intuitivamente», qué es bueno y qué es malo, si, por consiguiente, opinan que ya no tienen necesidad del cristianismo como garantía de la moral, eso mismo es simplemente *consecuencia* del dominio del juicio cristiano de valor y expresión de la *fortaleza* y *profundidad* de ese juicio: hasta el punto de que se ha olvidado el origen

> de la moral inglesa, hasta el punto de que no se percibe ya el carácter muy condicionado de su derecho a existir. Para el inglés la moral no es todavía un problema [...] (*Ibid.*: 70 [1980b: 88])

En este pasaje un tanto ambiguo, el objetivo de Nietzsche parece dar un giro hacia la «moralidad», la moralidad misma. Su tesis no es únicamente, pues, que hay ciertos elementos presentes en la moralidad cristiana que dejan de tener sentido una vez que se pierde la creencia en la existencia del Dios cristiano; más bien tiene que ser, según parece, la de que hay algunos rasgos generales de la moralidad cristiana que ha legado a todas las demás moralidades, que tienen que estar presentes en cualquier otro sistema de pensamiento y práctica para que ese sistema sea un sistema *moral*, y que *sólo tienen sentido dentro* del contexto de la creencia en la existencia del Dios cristiano. ¿Cuáles podrían ser estos rasgos generales? Un conjunto de candidatos serían los conceptos de obligación moral, de deber moral, de lo que es moralmente correcto o incorrecto y del sentido moral del término «debería». En efecto, una filósofa contemporánea nos dice que esos conceptos «deberían ser desechados si ello fuera psicológicamente posible», ya que se trata de «supervivencias, o derivaciones de supervivencias, de una concepción más antigua de la ética que en general ya no sobrevive» (Anscombe 1958: 1). Tampoco resulta difícil conjeturar cuál era esa «concepción más antigua de la ética»: la concepción que explicaba los conceptos citados en términos de los mandatos del Dios cristiano.

Repito: la tesis en cuestión no es que haya algunos elementos presentes en la moralidad cristiana que dejen de tener sentido una vez que se pierde la creencia en la existencia del Dios cristiano. Y acepto: es difícil entender de qué manera un sistema de pensamiento y práctica que desecha *todo* lo que Anscombe dice que debería desecharse podría continuar considerándose un sistema de pensamiento y práctica *morales*. Pero esa aceptación no equivale, sin embargo, a la aceptación de la tesis en cuestión: la tesis de que la idea de que existen hechos morales, expresados en términos de los conceptos mencionados, tiene sentido únicamente bajo el supuesto de que existen hechos relativos a lo que ordena el Dios cristiano o bien bajo un análogo derivado de esa suposición.

La moralidad cristiana no es la única moralidad; y la tesis de que cualquier otra moralidad que emplee la idea de que hay hechos relativos a las obligaciones y los deberes morales *tiene que* verse perseguida por el fantasma del Dios cristiano y sus mandatos, sólo puede adquirir atractivo gracias a la influencia de dos fallas «metodológicas». Una es el apego a

un tratamiento casi enteramente abstracto de ciertas nociones morales ciertamente generales (véase Strawson 1961: 33); la otra, reconocida por el propio Nietzsche (1956b: 209), consiste en la incapacidad de ver la estricta falta de pertinencia de las cuestiones relativas a los orígenes para las cuestiones de utilidad o propósito actuales. Una vez que se eluden ambas fallas, podemos lograr una comprensión no problemática de *ciertos* usos *actuales* de las nociones que Anscombe pretende que desechemos enteramente:

> No hay nada mínimamente misterioso o metafísico en el hecho de que los deberes y las obligaciones acompañen a los cargos, los puestos y las relaciones con los demás. Las exigencias que se hacen a una persona en virtud de que ocupa cierto puesto pueden en efecto precisarse, y suelen precisarse explícitamente con mucho detalle. Y cuando decimos que alguien es consciente o decimos que tiene un fuerte sentido de sus obligaciones o del deber, por lo general no queremos decir que lo persiga el fantasma de la idea de los ordenamientos sobrenaturales; sino más bien pretendemos decir algo como lo siguiente: que se puede contar con su esfuerzo sostenido para realizar lo que se requiere de él en ciertas tareas definidas, para cumplir lo que se exige de él como estudiante o maestro o padre o soldado o lo que él sea. (Strawson 1961: 33)

Exactamente. Y sospecho que las mismas fallas «metodológicas» que pueden cegarnos ante esas verdades, pueden también cegarnos ante las verdades acerca de una cuestión afín.

La concepción cristiana original de los hechos morales, expresada en términos de los mandatos del Dios cristiano, en términos de «ordenamientos sobrenaturales», se acompañaba, desde luego, de la creencia en la capacidad de los seres humanos para obedecer y desobedecer libremente esos mandatos. Y el concepto general del «libre albedrío» así introducido se encapsuló luego en las concepciones de ciertos vicios y virtudes característicamente cristianos. Ahora bien, en determinado momento —al escribir acerca de Schopenhauer—, Nietzsche observa, con respecto a «la "voluntad"», que es «el más grande fraude psicológico que, descontado el cristianismo, hay en la historia» (1968b: 80 [1980b: 100]). Se trata de un fraude porque postula «*causas* imaginarias» (1968a: 125 [1979b: 39]), porque introduce el «falso» concepto de «causalidad espiritual» (*ibid.*: 151 [1979b: 69]). Al proceder así, produce un choque inmediato entre cualquier sistema de pensamiento y práctica que incorpore este postulado y la ciencia: «*Moraleja*: la ciencia es lo prohibido en sí —ella es lo único prohi-

bido. La ciencia es el *primer* pecado, el germen de todo pecado, el pecado *original. La moral no es más que esto.*» (Nietzsche 1968a: 164 [1979b: 84]).

Como antes, el objetivo de Nietzsche parece dar un giro en la discusión sobre este tema; pero, con el riesgo de convertirnos en el peor de los lectores, podríamos intentar la reconstrucción siguiente de uno de los pensamientos que aquí expresa. La concepción cristiana original de los hechos morales como «ordenamientos sobrenaturales» suponía la existencia del «libre albedrío», de la «causalidad espiritual»: un presupuesto falso y anticientífico de la verdad de lo que se ha denominado «la metafísica oscura y timorata de los partidarios del libre albedrío» (Strawson 1962: 25). Y ese presupuesto ha sido legado a todo sistema de pensamiento y práctica *morales* digno, por así decirlo, del nombre. El contenido de ese presupuesto es un componente esencial de la teoría interna del pensamiento y la práctica morales; la falsedad de ese presupuesto representa «el problema de la moralidad misma».

Debemos hacer tres concesiones a ese punto de vista del asunto. En primer lugar, concedamos que Nietzsche tiene razón acerca del sitio que ocupa cierto concepto general del «libre albedrío» dentro de la concepción cristiana original de los hechos morales. En segundo lugar, admitamos también que hay ciertos elementos dentro de esa concepción original que presuponen la coherencia de un concepto oscuro y timorato de «causalidad espiritual». Y, por último, admitamos también que muchos adeptos a las moralidades manifiestan su apego a esta metafísica oscura y timorata de la doctrina del libre albedrío. La cuestión estriba ahora en que aceptar todo lo anterior no nos permite aceptar la conclusión que Nietzsche pretende extraer.

Para entender esto, recordemos primero la estricta falta de pertinencia de los asuntos relativos a los orígenes para las cuestiones de uso o propósito actual. Acto seguido, pasemos por alto todos los elementos del tipo al que se hace referencia en la segunda de las concesiones que hemos hecho. Y, finalmente, veamos ahora los posibles resultados de abandonar el nivel de consideración casi enteramente abstracto de este asunto para favorecer una consideración detallada de las prácticas de aquellos a quienes se hace referencia en la tercera de las concesiones que hemos hecho: una consideración de sus prácticas específicas, digamos, no una consideración de lo que dicen, en un nivel casi enteramente abstracto, acerca de esas prácticas. La siguiente posibilidad puede ser ahora reconocida incluso por los adeptos que manifiesten apego a la metafísica de la doctrina del libre albedrío: la de que en sus prácticas específicas —por ejemplo,

en los detalles de sus *usos* específicos de los conceptos de responsabilidad y de confiabilidad y otros conceptos más específicos relacionados con ellos—, el apego así expresado resulta *vano*. Es decir, la consideración de sus prácticas específicas revela que tal apego no desempeña ningún papel en la generación de esas prácticas. Este apego es un legado de la moralidad cristiana y quizá de la concepción kantiana de la moralidad, pero, no obstante, constituye un legado *externo* en relación con su práctica moral. Ésta es la posibilidad que supongo que Strawson exploró en su artículo clásico sobre el tema (Strawson 1962). Supongo que ha demostrado, además, que por lo menos muchas de *nuestras* prácticas relativas a las atribuciones de responsabilidad o de confiabilidad quedarían enteramente intactas si la ciencia estableciera la verdad de algunas tesis deterministas incompatibles con la doctrina del libre albedrío: «no se seguiría de esas tesis [deterministas] que nadie decide hacer nada, que nadie hace jamás algo intencionalmente, que es falso que a veces las personas sepan a la perfección lo que están haciendo» (*ibid.*: 3). Es decir, las atribuciones de responsabilidad y confiabilidad que efectivamente hacemos, se fundan en consideraciones conceptuales que son neutrales con respecto al debate abstracto entre «deterministas» y «partidarios del libre albedrío», de suerte que cualquier posición que se adopte dentro de los términos de ese debate es una posición externa a nuestra práctica efectiva de hacer tales atribuciones. Por lo tanto, no veo razón alguna para pensar que la «moralidad misma», en virtud de su uso de los conceptos de responsabilidad y confiabilidad, *requiere* una posición dentro de ese debate.

Pero permítaseme añadir algunos reparos. En primer lugar, de lo que se acaba de expresar no se desprende que se niegue la posibilidad de una interacción mutua entre las prácticas y las tesis de los participantes acerca de esas prácticas. Tampoco se desprende de lo expresado que las tesis de los participantes acerca de sus prácticas carezcan siempre de pertinencia cuando se considera la cuestión de las teorías internas de esas prácticas. Y, finalmente, no forma parte de la explicación presente afirmar que la posibilidad en cuestión —la de los compromisos «teóricos» de los participantes que son externas a sus prácticas reales— es una posibilidad que, al realizarse, siempre se detecta con facilidad. De ahí proviene una observación más reciente de Strawson:

> Hay una ambigüedad muy general en la noción de «nuestro concepto común» de lo que sea. ¿Tienen que extraerse los rasgos característicos de un concepto semejante exclusivamente de su uso, de nuestra *práctica* común, o debemos añadir los agregados de la reflexión, por confusos que sean, los

> cuales, natural o históricamente, se acumulan a su alrededor? La distinción no está nada clara, pero donde es posible establecerla, prefiero la primera opción. (Strawson 1980: 265)

Pero esta precaución indispensable deja intacta la cuestión principal. No parece haber una imposibilidad mayor en una moralidad libre de «*causas* imaginarias», libre de «causalidad espiritual», que en una libre del «fantasma de la idea de los ordenamientos sobrenaturales» —o que en una libre de la obsesión característica de la moral cristiana por el asunto de la humildad, la culpa y el castigo—. Y ello parece estar confirmado por la existencia real, difundida o no, de moralidades libres de semejante parafernalia. Todavía vale la pena apuntar al blanco en el que verdaderamente dan las críticas de Nietzsche; pero no hay que confundirlo con la «moralidad misma».

Alasdair MacIntyre sostiene que las críticas de Nietzsche a la moralidad descansan en una generalización ilegítima de una condición de la moralidad temporalmente particular y seriamente desordenada (MacIntyre 1981: 107). Pero si lo que acaba de exponerse es correcto, la verdad parece ser un tanto cuanto diferente. Algunas de las críticas de Nietzsche de la moralidad pueden entenderse, en efecto, como señalamientos apoyados en una generalización ilegítima —pero desde las teorías particulares internas de ciertas moralidades basadas en la teología—, mientras que otras parecen descansar más bien en el presupuesto de que es Kant quien nos ha proporcionado la descripción filosófica correcta de la teoría interna de la «moralidad misma». Ampliando el asunto, podríamos decir que estas últimas críticas descansan en una generalización ilegítima de las verdaderas moralidades kantianas. Pero creo que la siguiente constituye una observación más útil: las críticas de la moralidad consideradas hasta aquí se apoyan en una identificación equivocada de la teoría interna de la «moralidad misma», una identificación equivocada de la institución general de la moralidad.

6.8 Si dios ha muerto

Quiero ahora apartarme de las tentativas de Nietzsche por demostrar que la teoría interna de la moralidad es incoherente o falsa o carente de sentido para considerar algunos tipos de críticas distintos que pueden dirigirse contra la moralidad. El puente para hacer este movimiento no posee más que cuatro tablones, dos en mis propias palabras y dos en las de Nietzsche:

(1) en general, las moralidades son notablemente ineficaces en lo tocante a la influencia, «activa» y «personal», que tiene sobre sus adeptos: a la vez que facilitan demasiado bien los juicios morales negativos acerca de otros y sus acciones —un hecho que no sorprenderá a quienes comprendan la discusión de Nietzsche sobre la primacía conceptual del mal o mi exposición mandevilliana de la pragmática del discurso moral—, las moralidades son, en general, ineficaces en cuanto a la influencia moral que tienen sobre uno mismo, sobre el carácter y las motivaciones propias;
(2) las excepciones a esa primera observación nos las ofrecen en general las moralidades fanáticas y repugnantes o partes de ellas, las moralidades inmorales o partes de ellas;
(3) «nos es lícito sentar como tesis suprema que, para *hacer* moral, es preciso tener la voluntad incondicional de lo contrario» (Nietzsche 1968b: 58 [1980b: 75]); y
(4) «*todos* los medios con que se ha pretendido hasta ahora hacer moral a la humanidad han sido radicalmente *inmorales*» (*ibid.*: 59 [1980b: 75]).

Hay más de una diferencia temática entre las observaciones de Nietzsche y las mías. Nietzsche se interesa en la cuestión *moralizadora* de los medios que podrían adoptarse razonablemente con miras al propósito de volver más moral a un *grupo de otras personas* (quizá tan numeroso como la humanidad —menos uno mismo—). El pensamiento de Nietzsche sobre esa cuestión parece ser el siguiente: una vez que comprendamos que Dios no anda por ahí para arreglar las cosas de otro modo, podríamos descubrir que los únicos medios moralizadores que podrían funcionar de hecho para el fin fijado son en sí mismos inmorales. Como lo expresa Philippa Foot: «Si Dios está muerto nada garantiza que el mal no pueda ser la condición del bien» (Foot 1973: 87 [1994: 107]). En cuanto afirmación sobre lo que podría suceder, eso parece casi innegable, pero examinar si realmente ocurre así resultaría un asunto complicado. Tal examen tendría que distinguir por lo menos tres tipos de contextos: aquellos en los que los sujetos que han de ser moralizados carecen de los conceptos morales pertinentes; aquellos en los que los sujetos en cuestión, aunque en apariencia poseen esos conceptos, no ven ninguna conexión entre los juicios expresados en sus términos y las razones para actuar; y aquellos en los que, aunque los sujetos ven dicha relación, la fuerza de las razones para actuar resultantes es comparativamente débil. Sin embargo, aquí prefiero considerar el análogo de la cuestión moralizadora de Nietzsche en el caso

individual, a saber, la cuestión de los medios que podría razonablemente adoptar un agente *individual* con el propósito de volverse *él mismo* más moral. Pero mi preferencia por esta cuestión individual no significa negar la importancia de la cuestión colectiva «¿Cómo podríamos volvernos *nosotros mismos* más morales?»; menos aún significa negar la importancia del asunto relativo a por qué la cuestión individual se plantea con mucho menor frecuencia que la cuestión moralizadora de Nietzsche en el pensamiento moral real.

Por cierto, mi cuestión individual cubre todavía un terreno demasiado amplio; en lo que sigue restringiré sus límites un poco más. Anteriormente (p. 97) nos topamos con una discusión de Judith Baker sobre una propuesta que Grice ha considerado en relación con el debate «internalismo-externalismo». La propuesta era ésta: «si Juan piensa que debería ejecutar una acción *a*, entonces ello exige que o bien Juan quiera *a* o bien que piense que debería querer *a*». Esta propuesta «puede aplicarse de manera regresiva», lo cual lleva a Judith Baker a sugerir que para la «gente real» los juicios originales de obligación «se concretarán» en un deseo *en algún nivel*: así, siendo Juan una persona real, o bien querrá *a* o querrá querer *a* o querrá querer querer *a* o... Ahora bien, supongamos que, en un caso dado, el juicio de obligación «se manifiesta» en un deseo del segundo nivel: Juan piensa que debería ejecutar una acción *a*, de hecho no quiere hacer *a*, pero sí quiere querer hacer *a*. Entonces Juan sabe que las consideraciones morales no son suficientes para producir el deseo de actuar de primer nivel, sino únicamente el deseo de segundo nivel; pero dado que de hecho tiene ese deseo de segundo nivel, Juan podría buscar otras consideraciones, no morales, cuyo reconocimiento sería suficiente para producir en él el deseo de primer nivel faltante. Y si esa búsqueda tiene éxito, si encuentra esas consideraciones no morales y surge el correspondiente deseo de primer nivel, Juan podría actuar de manera tal que, externamente, su acción coincidiera con el contenido de su obligación reconocida.

Pero, ¿tendría esa acción valor moral? Kant sostenía que para la acción moral nuestros motivos tienen que ser puros: una acción sólo posee «genuino mérito moral» si se ejecuta «*únicamente* por deber y sin inclinación alguna» (1959: 14, las cursivas son mías). Es difícil interpretar ese enunciado: para eludir problemas de exégesis del texto, atribuyamos a Kant, con o sin justicia, la doctrina de que, para que una acción tenga valor moral, las *únicas* consideraciones motivadoras en favor de esa acción que el agente puede contemplar en sus deliberaciones prácticas

han de ser de índole moral. De esta concepción se sigue que, en el caso que acabamos de describir, la acción de Juan carece de valor moral. Pero aunque la conclusión parezca inteligible conforme a cierto modelo de agentes morales ideales, también podría parecer demasiado dura con respecto a la «gente real». Así, puede sostenerse la interpretación de que Judith Baker defiende una postura kantiana *modificada* que admite que la acción de Juan pueda en verdad tener un «genuino mérito moral». Pues desde el punto de vista de esta posición kantiana modificada, se admite la entrada, en las deliberaciones prácticas del agente, de otras consideraciones motivadoras, no morales, en favor de la acción de que se trate, siempre y cuando estén subordinadas a las consideraciones morales: la *primacía* de las consideraciones morales motivadoras se juzga compatible con cierto papel *secundario* o *derivado*, dentro de las deliberaciones prácticas del agente, para otras consideraciones motivadoras no morales. Y, al parecer, en esta posición kantiana modificada se admitiría una posibilidad similar para los casos en los cuales, aunque el agente posee en efecto un deseo pertinente del primer nivel, fundado únicamente en consideraciones morales, duda, sin embargo, de que la fuerza de ese deseo sea suficiente para desembocar en la acción correspondiente. Tiene un deseo de segundo nivel de que su deseo del primer nivel sea más fuerte: así pues, podrá buscar otras consideraciones motivadoras para *respaldar* el deseo moral.

El primer aspecto que debe advertirse acerca de esta posición kantiana modificada desemboca en una preocupación análoga a la queja de Nietzsche en el caso de los grupos de otras personas. Para que esa posición ofrezca una explicación convincente de las circunstancias en las cuales una acción tiene un «genuino mérito moral», es sin duda necesario que se incorpore en ella una condición adicional: que las consideraciones motivadoras secundarias *no sean inmorales*. Si esa motivación secundaria se fundara en la maldad o en la crueldad, entonces, a pesar de cualquier coincidencia externa con la obligación del agente, es indudable que su acción no podría tener un «genuino mérito moral». Más aún, esa condición podría no ser satisfecha por una amplia gama de casos. Supongamos, por ejemplo, que la motivación secundaria en cuestión está basada en el deseo de placer. Resultará necesario, antes de aceptar que la acción en cuestión tiene un «genuino mérito moral», considerar la posibilidad de que la generación del placer en cuestión dependa de consideraciones inmorales —los diversos tipos de consideraciones involucradas, por ejemplo, en la hipocresía o el orgullo o la autosatisfacción o la jactancia o la vanidad

o la mojigatería—. A la teoría que incorpora esta condición adicional la llamaremos teoría kantiana *modificada modificada.*

Los valores morales no son los únicos valores prácticos que adoptan los seres humanos; por lo tanto, dentro de los términos de la teoría modificada modificada, el «genuino mérito moral» de una acción es compatible con el hecho de que sea resultado de un proceso deliberativo en el cual el agente recurre a consideraciones evaluativas de otros tipos, no morales, con el fin de originar el deseo del primer nivel conforme al cual actuó. Pero deben advertirse tres cuestiones acerca de esto. En primer lugar, esa posibilidad será negada por cualquier concepción moral que considere inmorales todos los demás valores no morales. En segundo lugar, ciertas concepciones acerca de las consideraciones motivacionales *eficaces* también tendrán que negar esa posibilidad. Para Nietzsche, «el principio absoluto y fundamental de la conducta humana es la voluntad de poder» (Foot 1973: 94 [1994: 114]), y ése es, en consecuencia, el único principio que podría estar vigente en cualquier motivación secundaria eficaz. Pero supongamos que eso se acepta junto con el pensamiento de que esa voluntad de poder es inmoral: no habrá entonces, desde la perspectiva de la posición kantiana modificada modificada, ninguna posibilidad de motivaciones secundarias que sean a la vez eficaces y moralmente admisibles. El resultado de la imposibilidad sólo se presenta bajo la suposición de que la teoría de Nietzsche sobre *el* principio que determina *todo* comportamiento humano es verdadera; pero, como lo expresa Philippa Foot, «las teorías generales sobre los orígenes de la acción son trampas para los filósofos» (*ibid.* [1994: *ibid.*]). Por lo tanto, debemos destacar, en tercer lugar, que la sustitución de la teoría de Nietzsche acerca de la motivación humana por algunas teorías más cautelosas todavía sirve para dar lugar a cierto escepticismo acerca de la probabilidad de las acciones de «genuino mérito moral». Por ejemplo,

> los seres humanos son egoístas por naturaleza [...] [Eso] parece verdad según los hechos, cuandoquiera y dondequiera que los observemos, pese a un reducidísimo número de excepciones evidentes. Acerca del carácter de este egoísmo, la psicología moderna ha encontrado algo que decirnos. La psique es un individuo históricamente determinado que cuida de sí mismo implacablemente. En cierto modo, se parece a una máquina; para funcionar necesita fuentes de energía y está predispuesta a ciertas pautas de actividad. El área de su alabada libertad de elección no es comúnmente muy amplia. Uno de sus pasatiempos principales consiste en soñar despierta. Se muestra reacia a encarar realidades desagradables. Su conciencia no es normalmente

> un cristal transparente a través del cual contempla el mundo, sino una nube de ensueños más o menos fantásticos destinada a proteger a la psique del dolor. Constantemente busca consuelo, ya sea a través de un imaginario hinchamiento del ego o a través de ficciones de naturaleza teológica. Hasta sus actos de amor son, con mucha frecuencia, una afirmación del ego. Me parece que es muy probable que nos reconozcamos en esta descripción bastante deprimente. (Murdoch 1970: 78–79)

Así se expresa Iris Murdoch bajo la influencia de Freud; si hay algo que capta el equivalente mundano de la doctrina del pecado original es, sin duda, esta descripción del hombre. Y quien la acepte pensará seguramente que, incluso desde el punto de vista de la posición kantiana modificada modificada, los casos de acciones de «genuino mérito moral» serán raros: pues en general cualquier motivación secundaria *eficaz* tendrá que basarse en el aspecto «naturalmente egoísta» de los seres humanos.

Sobre esa posición debe advertirse ahora un segundo asunto que viene a complicar las cosas. Puede atribuirse un «genuino mérito moral» a una acción de acuerdo con esa posición incluso cuando la acción sea el resultado de una consideración motivadora secundaria, siempre que esta consideración motivadora no sea moralmente peor que neutral. Pero, ¿cómo ha de entenderse exactamente este discurso sobre consideraciones motivadoras secundarias, o subordinadas, o derivadas? No basta con interpretar tal discurso desde un punto de vista meramente histórico-causal. Volvamos a considerar al agente Juan. Éste piensa que debería ejecutar cierta acción *a*, de hecho no quiere hacer *a*, pero de hecho quiere querer hacer *a*; por lo tanto, se ve conducido a buscar otras consideraciones motivadoras, no peores que moralmente neutrales, cuyo reconocimiento bastaría, dadas las circunstancias, para producir un deseo de primer nivel con un contenido que en lo externo coincidiera con el del deseo del primer nivel ahora faltante. Supongamos que su búsqueda tiene éxito: ha llegado entonces a apreciar esas otras consideraciones motivadoras sólo porque reconoce su obligación moral. Pero supongamos que ese reconocimiento desaparece y que, pese a ello, el agente termina por actuar de acuerdo con el deseo de primer nivel que ahora posee: ¿podría ser éste un caso de acción de «genuino mérito moral»? Además, no es suficiente imponer simplemente la condición adicional de que el reconocimiento, por parte del agente, de su obligación moral, junto con el correspondiente deseo de segundo nivel, continúa existiendo: pues ello es compatible con la posibilidad de que la satisfacción de su obligación no sea sino un efecto colateral (previsible) de su acción basada en su deseo de primer nivel.

Parece necesaria alguna otra condición que capte la idea de que las consideraciones motivadoras secundarias *siguen siendo secundarias* aun cuando sean las que dan origen a la acción. Hay varias candidatas a tal condición, pero cualquiera que sea la preferida, la teoría final encara una dificultad más: es probable que el alcance de la aplicación de la noción consecuente de «genuino mérito moral» se encuentre de hecho muy restringido a la luz de la dificultad empírica de conciliar la fuerza y la vivacidad de las consideraciones motivadoras secundarias que se requieren para producir el deseo de primer nivel con su *persistente* condición de secundarias.

Quizás con excesiva ambición, la moraleja que quiero extraer de esta discusión consiste en que *si* aceptamos una variante menos dura pero todavía convincente de la teoría kantiana acerca de qué significa que una acción posea «genuino mérito moral», hay buenas razones para suponer que, en general, se presente una dificultad considerable con respecto a las probabilidades de que haya acciones de «gente real» que tengan ese mérito.

Sin duda, mucho de lo que se ha dicho para el caso individual se transfiere al caso de la moralización de un grupo de otras personas; también es indudablemente cierto que, en el segundo caso, entran en juego otras consideraciones que tal vez presten apoyo a la afirmación de Nietzsche de que «*todos* los medios con que se ha pretendido hasta ahora hacer moral a la humanidad han sido radicalmente *inmorales*» (1968b: 59 [1980b: 75]). Pensemos, por ejemplo, en la «sanción social» del castigo y de lo que se hace pasar por «educación moral»; o pensemos, en términos más abstractos, en la concepción maniquea de Mandeville acerca del Universo. Más aún, si la preocupación primaria de nuestro pensamiento moral viene a ser la de juzgar y moralizar a *otros*, descuidaremos la consideración moral de los medios para ese fin —de *nuestros* medios para ese fin—, y con tal descuido el riesgo hacia el que Nietzsche llama la atención aumentará, quizá casi hasta el límite.

6.9 El topo y el enano

Quiero, por último, pasar de la consideración de los medios para alcanzar la moralidad a la consideración del tema distinto, aunque relacionado, de los efectos de la moralidad («relacionado», porque los medios también tienen efectos). Pero primero observemos que cuando ahora hablamos de los efectos de la moralidad, es poco probable que el discurso trate de los efectos de las acciones de «genuino mérito moral»: estas rarezas han de tener, muy probablemente, pocos efectos. Más bien, como ocurre con

ciertas tesis de Mandeville, ese discurso debe entenderse, en general, en términos de los efectos del empleo general (no especificado) de categorías morales del pensamiento.

Como es predecible, hay otra vez una dificultad con respecto al blanco de los ataques de Nietzsche. A veces está claro que su blanco es un grupo de virtudes característicamente cristianas, más frecuentemente la piedad y la compasión:

> [C]omo si la humildad, la castidad, la pobreza, en una palabra, la *santidad*, no hubiesen causado hasta ahora a la vida un daño indeciblemente mayor que cualesquiera horrores y vicios ... El espíritu puro es la mentira pura... (Nietzsche 1968a: 119–120 [1979b: 33]; véase Hume sobre las «virtudes monacales», pp. 151–152.)

> Reteniendo las verdades, garabateando cosas con mano de necio, con un corazón chiflado, y echando numerosas mentirillas de compasión: — así he vivido yo siempre entre los hombres. (Nietzsche 1961: 204 [1979a: 260])

> A quien vive entre los buenos la compasión le enseña a mentir. La compasión vicia el aire a todas las almas libres. La estupidez de los buenos es, en efecto, insondable. (*Ibid.*: 204–205 [1979a: 260])

> Ay, ¿en qué lugar del mundo se han cometido tonterías mayores que entre los compasivos? ¿Y qué cosa en el mundo ha provocado más sufrimiento que las tonterías de los compasivos? (*Ibid.*: 249 [1979a: 319])

Pero en ocasiones parece claro que Nietzsche tiene en mente un blanco más extenso, un blanco equivalente a la «moralidad misma»:

> Abstenerse mutuamente de la ofensa, de la violencia, de la explotación: equiparar la propia voluntad a la del otro: en cierto sentido grosero esto puede llegar a ser una buena costumbre entre los individuos, cuando están dadas las condiciones para ello (a saber, la semejanza efectiva entre sus cantidades de fuerza y entre sus criterios de valor, y la homogeneidad de los mismos dentro de *un solo* cuerpo). Mas tan pronto como se quisiera extender ese principio e incluso considerarlo, en lo posible, como *principio fundamental de la sociedad*, tal principio se mostraría en seguida como lo que es: como voluntad de *negación* de la vida, como principio de disolución y de decadencia. (Nietzsche 1966a: 393 [1980a: 221])

> Camino a través de este pueblo y mantengo abiertos los ojos: se han vuelto *más pequeños* y se vuelven cada vez más pequeños — *y esto se debe a su doctrina acerca de la felicidad y la virtud.*

> En efecto, también en la virtud son modestos – pues quieren comodidad. Pero con la comodidad no se aviene más que la virtud modesta. (Nietzsche 1961: 189 [1979a: 239])

Finalmente, uno de los pasajes más claros en que Nietzsche reúne la causa y el efecto:

> Cuando los instintos más elevados y más fuertes, irrumpiendo apasionadamente, arrastran al individuo más allá y por encima del término medio y de la hondonada de la conciencia gregaria, entonces el sentimiento de la propia dignidad de la comunidad se derrumba, y su fe en sí misma, su espina dorsal, por así decirlo, se hace pedazos: en consecuencia, a lo que más se estigmatizará y se calumniará será cabalmente a tales instintos. La espiritualidad elevada e independiente, la voluntad de estar solo, la gran razón son ya sentidas como peligro; todo lo que eleva al individuo por encima del rebaño e infunde temor al prójimo es calificado, a partir de este momento, de *malvado (böse)*; los sentimientos equitativos, modestos, sumisos, igualitaristas, la *mediocridad* de los apetitos, alcanzan ahora nombres y honores morales. (Nietzsche 1966a: 303–304 [1980a: 181–182])

Así, la moralidad expresa una expectativa y una exigencia «absurdas»: que «la fuerza *no* sea un querer-dominar, un querer-sojuzgar, un querer-enseñorearse» (Nietzsche 1956b: 178 [1981: 51]). Y el efecto de que así ocurra es una «forma de decadencia, esto es, de empequeñecimiento, del hombre, como su mediocrización y como su rebajamiento de valor» (1966a: 307 [1980a: 135]). A la luz de este hecho, necesitamos volver a la cuestión de «qué tipo de hombre se debe *criar*, se debe *querer*, como tipo más valioso, más digno de vivir, más seguro de futuro» (1968a: 116 [1979b: 28]); y si nos remitimos a la propia respuesta de Nietzsche a esa pregunta –si nos remitimos a la propagación de los precursores del *Übermensch*, a la producción de «un preludio de actores mejores»— nos daremos cuenta de que ello requiere el total abandono de la moralidad, de «toda moralidad», de la «moralidad misma».

Varias de las cuestiones que se suscitan en relación con esta nueva crítica de la moralidad y la moraleja alcanzada se han discutido ampliamente en otros lugares, de modo que aquí sólo tengo que mencionarlas brevemente. Una de las cuestiones que se suscitan está relacionada con el *tipo* de juicio de valor involucrado en el discurso de Nietzsche de la «decadencia» y del tipo de ser humano que posee más valor. Strawson podría tal vez sostener que tales juicios pertenecen a la región más general de la ética, en contraste con el área más específica de la moral: es decir, pertenecen a

la región de «las evaluaciones que *pueden* regir las elecciones y decisiones que son de la mayor importancia para los hombres», la región en la que se encuentran las tentativas de presentar «cierta imagen ideal de una forma de vida [...] capaz de producir una respuesta de la más viva simpatía por parte de aquellos cuyos estilos de vida propios no pueden estar más alejados de ajustarse a la imagen expresada» (Strawson 1961: 27-28). Philippa Foot considera el asunto con más detalle (Foot 1973: 90-93 [1994: 110-113]). Por un lado, el sistema de Nietzsche y el de la «moralidad tradicional y particularmente griega» tienen un terreno común en virtud de la insistencia de Nietzsche en la importancia de la autodisciplina, y, también, de su interés por la cuestión de «la forma en que deben vivir los hombres para *vivir bien*». (Aunque debemos advertir que gran parte de la «moralidad griega» parece pertenecer a la región más general que Strawson llama «la ética», y no al área de «la moral».) Pero, por otro lado, (i) en cierto momento, Nietzsche asimila explícitamente sus juicios de valor a los de la estética; (ii) rechaza cualquier idea de normas de conducta *específica* que sean uniformes en toda la comunidad; (iii) rechaza incluso la idea de que su tipo de virtud sea predicada a *todos* los hombres, y (iv) en apariencia, está dispuesto a abandonar las consideraciones de justicia con el fin de allanarle el camino al *Übermensch*. Así, aunque no hay una respuesta nítida a la pregunta, la balanza está claramente en favor de negarle el carácter moral a las evaluaciones de Nietzsche. Tal vez, por así decirlo, deberíamos considerarlo un *inmoralista ético*.

Una segunda cuestión se relaciona con el *contenido* de los juicios de valor de Nietzsche sobre el *Übermensch* (y por tanto sobre el prólogo a él y sobre los efectos de la moralidad). Más de un filósofo ha subrayado «la vaciedad del ideal nietzscheano» del *Übermensch* (MacIntyre 1966: 225), la ausencia de valores *determinados* en ese supuesto ideal: «La concepción de Nietzsche es puro *wind-up* sin lanzamiento», si se quiere (Nozick 1981: 566-567 n). Se puede ser *casi todo* de una manera creativa, autodisciplinada y amante de la vida. (Lo máximo que Russell pudo hacer con el ideal fue lo siguiente: «muy parecido a Sigfrido, salvo porque sabe griego»; y luego una digresión característica: «Esto puede parecer extraño, pero no es mi culpa» (Russell 1946: 788).) Pero dado que, según observó Alasdair MacIntyre (1966: 225), ese defecto fue uno de los rasgos de los escritos de Nietzsche que permitieron que fueran tergiversados a manos de su hermana, bien podríamos sentir que por él ha sido castigado póstumamente de un modo más que suficiente.

Una tercera cuestión, subrayada por Foot, es la cuestión *empírica* de si Nietzsche tiene razón en sus tesis referentes a los efectos históricos de la

moralidad y los efectos futuros de la puesta en práctica de su programa para mejores actores. Pero esa cuestión exigiría un afinamiento considerable antes de quese le pudiera dar alguna respuesta razonable: *La Moralidad* versus *La Vida* no es el mejor definido de los frentes de batalla.

Para terminar, sin embargo, quiero considerar otra cuestión: ¿cómo *podría* la moralidad tener los asombrosos efectos que Nietzsche le atribuye? ¿Cómo podría la moralidad, el sistema de meras *ideas* morales, ser históricamente tan importante y desde el punto de vista humano tan desastrosa? (Compárense las afirmaciones generalmente míticas acerca de la importancia que tiene la inmoralidad —las *acciones* inmorales— para la sociedad, la especie y el mundo.) Es cierto que Nietzsche piensa que la moralidad se ha creado precisamente para que tenga esos efectos; pero hemos de recordar la estricta falta de pertinencia de los asuntos relativos al origen para las cuestiones de uso o propósito —o efectos— actuales. Así pues, ¿qué características de la moralidad *tal como ahora existe* podrían darle poderes tan asombrosos? Cualquier respuesta que Nietzsche nos ofreciera sería una respuesta compleja; no obstante, creo que una gran parte del peso explicativo de esa respuesta recaería en las consideraciones siguientes:

> Mas a sí mismo se ha descubierto quien dice: éste es *mi* bien y éste es *mi* mal: con ello ha hecho callar al topo y enano que dice: «bueno para *todos*, malvado para *todos*». (Nietzsche 1961: 212 [1979a: 271])

> «Éste — es *mi* camino, — ¿dónde está el vuestro?», así respondía yo a quienes me preguntaban «por el camino». ¡*El* camino, en efecto, — no existe! (*Ibid.*: 213 [1979a: 272])

> La realidad nos muestra una riqueza fascinante de tipos, la exuberancia propia de un pródigo juego y mudanza de formas: ¿y cualquier pobre mozo de esquina de moralista dice a esto: «¡no!, el hombre debería ser *de otro modo*»? [...] Él sabe incluso *cómo* debería ser él, ese mentecato y mojigato, se pinta a sí mismo en la pared y dice *¡ecce homo!* (Nietzsche 1968b: 46 [1980b: 58])

> [H]ace la gente en su interior una moral, una virtud, una santidad, establece una conexión entre la *buena* conciencia y el ver las cosas de manera *falsa*, — exige que ninguna *otra* especie de óptica tenga ya valor... (Nietzsche 1968a: 120 [1979b: 33])

Ahora bien, aceptemos que es inherente a la «moralidad misma» el hecho de que por lo menos algunos valores morales sean, en cierto sentido, «universales» —«bueno para *todos*, malvado para *todos*»—, e incluso que esos valores son, en algún sentido, «sólidos» —«*Sobre* la corriente todo es

sólido, todos los valores de las cosas, los puentes, conceptos, todo el "bien" y el "mal": ¡todo eso es *sólido*!» (Nietzsche 1961: 218 [1979a: 279])—. Esas concesiones deben entenderse en términos que las vuelvan compatibles con nuestras discusiones anteriores sobre la subjetividad de los valores, las discrepancias morales y los «relativismos morales» (capítulos 3 y 5). Pero así entendidas, esas concesiones están lejos de ser suficientes para explicar cómo podría tener la moralidad los asombrosos efectos que Nietzsche le atribuye. Si fuera la moralidad *misma* lo que explica la «decadencia» del hombre, lo que explica la tendencia *descendente* hacia la *uniformidad*, la moralidad tendría que contener también en su seno la doctrina de que los valores morales son los «de más peso», si no la doctrina más fuerte de que los valores morales son los únicos valores verdaderos. Una vez más, Nietzsche parece pensar que fue Kant quien descubrió la teoría interna de la «moralidad misma»:

> [Kant] inventó una razón expresamente para averiguar en qué caso no hemos de preocuparnos por la razón, a saber, cuando la moral, la sublime exigencia del «tú debes» deja oír su voz. Si se tiene en cuenta que casi en todos los pueblos el filósofo no es más que el desarrollo ulterior del tipo sacerdotal, no sorprenderá ya esa parte de herencia del sacerdote, la *superchería frente a sí mismo*. Cuando uno tiene tareas sagradas, como, por ejemplo, las de mejorar, salvar, redimir a los hombres, cuando uno lleva en su pecho la divinidad, cuando es vocero de imperativos trasmundanos, está ya, con tal misión, fuera de todas las valoraciones meramente ajustadas al entendimiento... (Nietzsche 1968a: 123 [1979b: 36-37])

Pero si Nietzsche tuviera razón en esto, tanto la discusión de Strawson sobre las complejas relaciones entre las moralidades sociales y los ideales individuales, como mi discusión previa sobre las cuestiones referentes al alcance de la moralidad, habrían sido incoherentes; pero éstas no lo fueron, por lo tanto Nietzsche no tiene razón. Cualesquiera que sean las verdades históricas acerca de los acompañantes de la moralidad y sus efectos, esas verdades tienen que distinguirse de las verdades acerca de la naturaleza de la moralidad; y el encomiable deseo de censurar a algunos de esos acompañantes históricos tiene que separarse del deseo de criticar la «moralidad misma».

En lugar de preocuparme por el riesgo de figurar entre los peores lectores, me podría haber contentado quizá con hacer referencia a la observación de Nietzsche acerca de la historia de la filosofía con la que se inició la segunda parte de este ensayo (omitiendo —por favor— la palabra «mejor»),

Pero aunque he intentado no mancillar, ni confundir ni blasfemar, no creo que esta discusión sobre Nietzsche haya construido más de lo que ha desmantelado y dispersado. En particular, he tratado de mostrar que no hay una buena razón para creer que la teoría interna del pensamiento y la práctica morales sólo tiene sentido si se basa en presupuestos de carácter teológico (pp. 230–232), ni para creer que esa teoría contiene la metafísica oscura y timorata de los partidarios del libre albedrío (pp. 232–235), ni para creer que incluye alguna doctrina objetablemente fuerte de la soberanía de la moral sobre todos los demás valores prácticos (pp. 246–247). En cada uno de los casos surge el riesgo, señalado en la introducción de este ensayo, de que un escepticismo acerca de la moralidad se base en una falsa identificación de la teoría interna del pensamiento y la práctica morales. Se ha señalado aquí parte de lo que sí contiene esa teoría (p. 206; véanse también las pp. 231 y 243–246), y se han criticado en efecto algunos de los importantes escepticismos sobre la moralidad consecuentes (capítulos 4 y 5). Pero debo admitir que la concepción parcial de la moralidad que así surge no tiene probabilidades de suscitar —desde un punto de vista teórico o práctico, en sus adeptos o en sus críticos— el *Entusiasmo*, cuyas verdaderas fuentes son «la esperanza, el orgullo, la presunción, una ardiente imaginación, junto con la ignorancia» (Hume 1963b: 76); pero creo que eso no es tan malo. La evangelización, por muy meritoria que sea su causa, sigue siendo desagradable.

Sin embargo, en este capítulo final también he intentado indicar algunos rasgos de nuestro pensamiento y nuestra práctica morales que podrían constituir el punto de partida para una duda distintiva acerca de la moralidad. La combinación de la teoría de la naturaleza humana de Mandeville o de Nietzsche con la teoría kantiana de lo que significa que una acción posea un genuino mérito moral, produce la consecuencia de que tal acción sea imposible para los seres humanos (pp. 217–219 y 237–240). Pero incluso si sólo se combina una teoría más cautelosa, aunque todavía moralmente pesimista, de la naturaleza humana, con una concepción kantiana más débil pero convincente acerca del genuino mérito moral, se produce la consecuencia de que en general es poco probable que haya acciones humanas reales que tengan ese mérito (pp. 237–241). Esa concepción kantiana representa una respuesta posible a la pregunta, expresada en los términos de la primera parte de este ensayo, de si, y cuándo, las valoraciones morales dentro de las áreas conformadas por los deseos potenciales de los tipos (3) y (4) admiten, desde el punto de vista moral, el respaldo de otros tipos de valoraciones y deseos (capítulo 3); pero tengo que con-

fesar que estoy desesperadamente confundido respecto de si esa u otra concepción kantiana de lo que significa que una acción posea un genuino mérito moral forma parte realmente de la teoría interna del pensamiento y la práctica morales (véase la p. 219). Pero aun si sólo ocurriera que eso es lo que generalmente *se cree*, esa creencia podría combinarse con la aceptación general de una teoría moralmente pesimista de la naturaleza humana para producir ciertos rasgos *moralmente* desagradables del uso del pensamiento y la práctica morales, rasgos que se encuentran, me parece, en nuestro uso real (pp. 224–227). Quizá el reconocimiento de esos rasgos sea parcialmente responsable del interés reciente por la cuestión del «relativismo moral»; pero si esto es así, las propuestas «relativistas» consecuentes que restringen la crítica moral incurren en el error de presentar soluciones teóricas a un problema *práctico* —el mismo error que se encuentra en muchas discusiones sobre el alcance interno de la moralidad (pp. 169–170)—. Repito: no sé si haya alguna concepción kantiana sobre el genuino mérito moral de las acciones que forme parte de la teoría interna del pensamiento y la práctica morales. Pero si es así, y si en consecuencia las aspiraciones a la moralidad no se ajustan bien, en general, a nuestra naturaleza, a «nuestra constitución interior», entonces se requiere una *decisión*; y en este contexto, la réplica de que la idealización no es algo que haya que disculpar dentro de una teoría de la naturaleza de la acción moral, tendría que contraponerse por lo menos a una evaluación moral de los efectos reales de ese desajuste. Esto es, podría incluso resultar que las consideraciones morales sugirieran que la mayoría de nosotros debería abandonar los términos de referencia mismos de la institución de la moralidad.

POSFACIO

Había en Sevilla un loco que dio en el más gracioso disparate y tema que dio loco en el mundo. Y fue que hizo un cañuto de caña puntiagudo en el fin, y en cogiendo algún perro en la calle, o en cualquiera otra parte, con el un pie le cogía el suyo, y el otro le alzaba con la mano, y como mejor podía le acomodaba el cañuto en la parte que, soplándole, le ponía redondo como una pelota, y, en teniéndole de esta suerte, le daba dos palmaditas en la barriga, y le soltaba, diciendo a los circunstantes, que siempre eran muchos:

— ¿Pensarán vuesas mercedes ahora que es poco trabajo hinchar un perro? — ¿Pensará vuesa merced ahora que es poco trabajo hacer un libro?

MIGUEL DE CERVANTES Y SAAVEDRA

Los hombres de letras que trabajan bajo el influjo de este desasosiego experimentan síntomas completamente diferentes. Cuando envidian a una persona por sus prendas y erudición, su principal cuidado es ocultar hábilidosamente esta debilidad, empezando, generalmente, por negar y despreciar las buenas cualidades que envidian; escudriñan con cuidado las obras de sus víctimas, encuentran defecto en los mejores pasajes, no buscan más que los errores y no desean otra cosa que tener la ocasión de poder regalarse con una equivocación evidente; en sus censuras son tan capciosos como severos, hacen montañas de toperas y no sólo no perdonan ni la sombra de una falta, sino que exageran la más insignificante omisión convirtiéndola en el mayor de los desatinos.

BERNARD DE MANDEVILLE

BIBLIOGRAFÍA

Anscombe, G.E.M. (1958), «Modern Moral Philosophy», *Philosophy*, 33.

——— (1963), *Intention*, Oxford, Basil Blackwell.

——— (1991), *Intención*, trad. Ana Isabel Stellino, Barcelona, Paidós/Universidad Autónoma de Barcelona.

Baker, Judith (1986), «Do One's Motives Have to Be Pure?», en Grandy y Warner (comps.) (1986).

Blackburn, Simon (1984), *Spreading the Word*, Oxford, Clarendon, 1978, sobre el cual se hacen las referencias de página.

Brentano, Franz (1969), *The Origin of Our Knowledge of Right and Wrong*, trad. Roderick M. Chisholm y Elizabeth H. Schneewind, Londres Routledge and Kegan Paul.

Davidson, Donald (1963), «Actions, Reasons and Causes», *Journal of Philosophy*, 60.

——— (1967), «Causal Relations», *Journal of Philosophy*, 64; reimpreso en Davidson (1980), sobre el cual se hacen las referencias de página.

——— (1969), «How is Weakness of the Will Possible?», en Joel Feinberg (comp.), *Moral Concepts*, Oxford; reimpreso en Davidson (1980), sobre el cual se hacen las referencias de página.

——— (1976), «Hume's Cognitive Theory of Pride», *Journal of Philosophy*, 73; reimpreso en Davidson (1980), sobre el cual se hacen las referencias de página.

——— (1980), *Essays on Actions and Events*, Oxford, Clarendon.

——— (1994), *Ensayos sobre acciones y sucesos*, trad. Olbeth Hansberg, José Antonio Robles y Margarita M. Valdés, México, Instituto de Investigaciones Filosóficas-UNAM/Crítica.

Dent, N.J.H. (1984), *The Moral Psychology of the Virtues*, Cambridge, Cambridge University Press.

Elster, Jon (1984), *Ulysses and the Sirens: Studies in Rationality and Irrationality*, Cambridge, Cambridge University Press.

——— (1989), *Ulises y las sirenas. Estudios sobre racionalidad e irracionalidad*, trad. Juan José Utrilla, México, Fondo de Cultura Económica.

Falk, W.D. (1947-1948), «"Ought" and Motivation», *Proceedings of the Aristotelian Society*, 48; reimpreso en Sellars y Hospers (comps.) (1952), sobre el cual se hacen las referencias de página.

Foot, Philippa (1963), «Hume on Moral Judgement», en D.F. Pears (comp.), *David Hume: A Symposium*, Londres, Macmillan, 1963.

——— (1972a), «Morality as a System of Hypothetical Imperatives», *Philosophical Review*, 81.

——— (1972b), «Reasons for Actions and Desires», *Proceedings of the Aristotelian Society, Supplementary Volume*; reimpreso en Foot (1978b), sobre el cual se hacen las referencias de página.

——— (1973), «Nietzsche: the Revaluation of Values», en R.C. Solomon (comp.), *Nietzsche, A Collection of Critical Essays*, Nueva York; reimpreso en Foot (1978b), sobre el cual se hacen las referencias de página.

——— (1978a), «Moral Relativism», The Lindley Lecture at the University of Kansas; reimpreso en Meiland y Krausz (comps.) (1982), sobre el cual se hacen las referencias de página.

——— (1978b), *Virtues and Vices, and Other Essays in Moral Philosophy*, Oxford, Blackwell.

——— (1994), *Las virtudes y los vicios*, trad. Claudia Martínez, México, Instituto de Investigaciones Filosóficas-UNAM (Filosofía Contemporánea).

Gallie, W.B. (1955-1956), «Essentially Contested Concepts», *Proceedings of the Aristotelian Society*, 56.

Genet, Jean (1963), *Our Lady of the Flowers*, trad. Bernard Frechtman, Nueva York, Grove Press.

Grandy, R.E. y R. Warner (comps.) (1986), *Philosophical Grounds of Rationality: Intentions, Categories, Ends*, Oxford, Clarendon.

Hampshire, Stuart (1959), *Thought and Action*, Londres, Chatto and Windus.

Harman, Gilbert (1975), «Moral Relativism Defended», *Philosophical Review*, 84; reimpreso en Meiland y Krausz (comps.) (1982), sobre el cual se hacen las referencias de página.

Honderich, Ted (comp.) (1985), *Morality and Objectivity*, Londres, Routledge and Kegan Paul.

Hume, David (1888), *A Treatise of Human Nature*, ed. e índice analítico de L.A. Selby-Bigge, Oxford, Oxford University Press.

——— (1902), *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, ed. L.A. Selby-Bigge, Oxford, Oxford University Press.

——— (1945a), *Investigación sobre el entendimiento humano*, trad. Juan Adolfo Vázquez, Buenos Aires, Losada (Biblioteca Filosófica).

——— (1945b), *Investigación sobre la moral*, trad. Juan Adolfo Vázquez, Buenos Aires, Losada (Biblioteca Filosófica).

——— (1963a), «Of Luxury», en David Hume, *Essays and Treatises on Several Subjects*, Oxford, Oxford University Press.

——— (1963b), «Of Superstition and Enthusiasm», en David Hume, *Essays: Moral, Political and Literary*, Oxford, Oxford University Press.

——— (1977), *Tratado de la naturaleza humana*, trad. Félix Duque, Madrid, Editora Nacional, 1977.

Kant, Immanuel (1959), *Foundations of the Metaphysics of Morals*, trad. Lewis White Beck, Nueva York, Liberal Arts Press.

Kenny, Anthony (1963), *Action, Emotion and Will*, Londres, Routledge and Kegan Paul.

Law, William (1845), *Remarks upon a Late Book Entitled The Fable of the Bees*, Londres, Routledge/Thoemmes Press.

MacIntyre, Alasdair (1966), *A Short History of Ethics*, Londres, Routledge.

——— (1981), *After Virtue*, Londres, Duckworth.

Mackie, J.L. (1977), *Ethics: Inventing Right and Wrong*, Harmondsworth, Penguin Books.

——— (1980), *Hume's Moral Theory*, Londres, Routledge and Kegan Paul.

Mandeville, Bernard de (1711), *A Treatise of the Hypochondriack and Hysteric Passions, Vulgarly Call'd the Hypo in Men and Vapours in Women... in Three Dialogues*, Londres, J. Tonson.

——— (1720), *Free Thoughts on Religion, the Church, and National Happiness*, Londres, T. Jauncy and J. Roberts.

——— (1724), *A Modest Defence of Publick Stews: or, An Essay upon Whoring as it Is Now Practis'd in These Kingdoms*, Londres, A. Moore.

——— (1732a), *A Letter to Dion, Occasion'd by his Book Call'd Alciphron, or The Minute Philosopher*, Nueva York, Arno Press.

——— (1732b), *An Enquiry into the Origin of Honour, and the Usefulness of Christianity in War*, Londres, J. Brotherton.

——— (1924), *The Fable of the Bees, Part II* (comp. F.B. Kaye), Indianapolis, Liberty Classics.

——— (1970), *The Fable of the Bees*, comp. Phillip Harth, Harmondsworth, Penguin.

——— (1982), *La fábula de las abejas, o Los vicios privados hacen la prosperidad pública*, comentario crítico, histórico y explicativo de F.B. Kaye, trad. José Ferrater Mora, México, Fondo de Cultura Económica.

McDowell, John (1978), «Are Moral Requirements Hypothetical Imperatives?», *Proceedings of the Aristotelian Society, Supplementary Volume*.

——— (1979), «Virtue and Reason», *The Monist*, 62.

——— (1982), «Reason and Action», *Philosophical Investigations*, 5.

——— (1985), «Values and Secondary Qualities», en T. Honderich (comp.) (1985).

McGinn, Colin (1983), *The Subjective View. Secondary Qualities and Indexical Thoughts*, Oxford, Clarendon.

Meiland, Jack. W y Michael Krausz (comps.) (1982), *Relativism: Cognitive and Moral*, Notre Dame, Indiana, University of Notre Dame.

Mill, John Stuart (1910), *Utilitarianism, Liberty, Representative Government*, Londres, J.M. Dent.

Monro, Hector (1975), *The Ambivalence of Bernard Mandeville,* Oxford, Clarendon.

Murdoch, Iris (1970), *The Sovereignty of Good,* Londres, Routledge and Kegan Paul.

Nagel, Thomas (1970), *The Possibility of Altruism,* Oxford, Clarendon.

——— (1978), «The Justification of Equality», *Crítica. Revista Hispanoamericana de Filosofía,* vol. X, no. 28; reimpreso en Nagel (1979), sobre el cual se hacen las referencias de página.

——— (1979), *Mortal Questions,* Cambridge, Cambridge University Press.

Nietzsche, Friedrich (1956), (a) *The Birth of Tragedy* y (b) *The Genealogy of Morals,* trad. Francis Golffing, Nueva York, Doubleday.

——— (1961), *Thus Spoke Zarathustra,* trad. R.J. Hollingdale, Harmondsworth, Penguin Books.

——— (1966a), *Beyond Good and Evil,* en Walter Kaufman (trad. y comp.), *Basic Writings of Nietzsche,* Nueva York, Modern Library.

——— (1966b), *Mixed Opinions and Maxims,* en Walter Kaufman (trad. y comp.), *Basic Writings of Nietzsche,* Nueva York, Modern Library.

——— (1966c), *The Wanderer and his Shadow,* en Walter Kaufman (trad. y comp.), *Basic Writings of Nietzsche,* Nueva York, Modern Library.

——— (1968) (a) *The Anti-Christ* y (b) *Twilight of the Idols,* trad. R.J. Hollingdale, Harmondsworth, Penguin Books.

——— (1979a), *Así habló Zaratustra,* 7a. ed., introd., trad. y notas de Andrés Sánchez Pascual, Madrid, Alianza Editorial (El libro de bolsillo, 377).

——— (1979b), *El Anticristo,* 6a. ed., introd., trad. y notas de Andrés Sánchez Pascual, Madrid, Alianza Editorial (El libro de bolsillo, 507).

——— (1980a), *Más allá del bien y del mal. Preludio de una filosofía del futuro,* 6a. ed., introd., trad. y notas de Andrés Sánchez Pascual, Madrid, Alianza Editorial, (El libro de bolsillo, 406).

——— (1980b), *Crepúsculo de los ídolos,* 4a. ed., introd., trad. y notas de Andrés Sánchez Pascual, Madrid, Alianza Editorial (El libro de bolsillo, 467).

——— (1981), *La genealogía de la moral,* 6a. ed., introd., trad. y notas de Andrés Sánchez Pascual, Madrid, Alianza Editorial (El libro de bolsillo, 356).

Nozick, Robert (1981), *Philosophical Explanations,* Cambridge, Mass., Harvard University Press.

O'Shaughnessy, Brian (1980), *The Will: A Dual Aspect Theory,* Cambridge, Cambridge University Press.

Pears, David (1964), «Predicting and Deciding», *Proceedings of the British Academy,* 50.

Platts, Mark (1979), *Ways of Meaning. An Introduction to a Philosophy or Lenguage,* Londres, Routledge and Kegan Paul.

——— (1980a), «Kind Words and Understanding», *Crítica. Revista Hispanoamericana de Filosofía,* vol. XII, no. 36, pp. 3–38 .

——— (1980b), «Moral Reality and the End of Desire», en Mark Platts (comp.) (1980c).

——— (1980c), *Reference, Truth and Reality. Essays on the Philosophy of Language*, Londres, Routledge and Kegan Paul.
——— (1983a), «Explanatory Kinds», *British Journal for the Philosophy of Science*, 34.
——— (1983b), «La naturaleza del mundo moral», *Análisis Filosófico*, 3.
——— (1992), *Sendas del significado*, trad. Cecilia Hidalgo y Eduardo A. Rabossi, revisión de Manuela Rius en colaboración con Antonio Zirión, México, Instituto de Investigaciones Filosóficas-UNAM/Fondo de Cultura Económica.

Putnam, Hilary (1981), *Reason, Truth and History*, Cambridge, Cambridge University Press.

Reid, Thomas (1969), *Essays on the Active Powers of Man*, Cambridge, Mass., The MIT Press.

Richards, David A.J. (1971), *A Theory of Reasons for Action*, Oxford, Clarendon.

Russell, Bertrand (1921), *The Analysis of Mind*, Londres, George Allen and Unwin.
——— (1946), *A History of Western Philosophy*, Londres, George Allen and Unwin.
——— (1962) *Análisis del espíritu*, 3a. ed., trad. Eduardo Prieto, Buenos Aires, Paidós.

Schiffer, Stephen (1976), «The Paradox of Desire», *American Philosophical Quarterly*, 13.

Sellars, Wilfrid y John Hospers (comps.) (1952), *Readings in Ethical Theory*, Nueva York, Appleton-Century-Crofts.

Smith, Michael (1987), «The Humean Theory of Motivation», *Mind*, 96, Oxford, Mind Association.

Strawson, P.F. (1959), *Individuals. An Essay in Descriptive Metaphysics*, Londres, Methuen.
——— (1961), «Social Morality and Individual Ideal», *Philosophy*, 36; reimpreso en Strawson (1974), sobre el cual se hacen las referencias de página.
——— (1962), «Freedom and Resentment», *Proceedings of the British Academy*, 48; reimpreso en Strawson (1974), sobre el cual se hacen las referencias de página.
——— (1974), *Freedom and Resentment and Other Essays*, Londres, Methuen.
——— (1980), «Reply to Ayer and Bennett», en Zak van Straaten (comp.) (1980).
——— (1985), *Skepticism And Naturalism: Some Varieties*, Londres, Methuen.
——— (1989), *Individuos*, trad. Alfonso García Suárez y Luis Valdés Villanueva, Madrid, Taurus.

Stroud, Barry (1977), *Hume*, Londres, Routledge and Kegan Paul.
——— (1995), *Hume*, 2a. ed., trad. de Antonio Zirión, México, Instituto de Investigaciones Filosóficas-UNAM.

Taylor, Richard (1984), *Good and Evil*, Buffalo, N.Y., Prometheus.

Urmson, J.O. (1946), «On Grading», *Proceedings of the Aristotelian Society, Supplementary Volume*.
——— (1958), «Saints and Heroes», en A.I. Melden (comp.), *Essays in Moral Philosophy*, Seattle, University of Washington.

Warnock, G.J. (1971), *The Object of Morality,* Londres, Methuen.

Wiggins, David (1976), «Truth, Invention and the Meaning of Life», *Proceedings of the British Academy,* 62; reimpreso en Wiggins (1987b), sobre el cual se hacen las referencias de página.

——— (1980), *Sameness and Substance,* Oxford, Basil Blackwell.

——— (1985), «Claims of Need», en Ted Honderich (comp.) (1985).

——— (1987a), «A Sensible Subjectivism?», en David Wiggins (1987b).

——— (1987b), *Needs, Values, Truth. Essays in the Philosophy of Value,* Oxford, Basil Blackwell.

Williams, Bernard (1972), *Morality: An Introduction to Ethics,* Harmondsworth, Penguin Books.

——— (1973), «Egoism and Altruism», en Bernard Williams, *Problems of the Self,* Cambridge, Cambridge University Press.

——— (1974-1975), «The Truth in Relativism», *Proceedings of the Aristotelian Society,* 75.

——— (1981a), «Internal and external reasons», en Williams (1981c).

——— (1981b), «Practical Necessity», en Williams (1981c).

——— (1981c), *Moral Luck,* Cambridge, Cambridge University Press.

——— (1985), *Ethics and the Limits of Philosophy,* Londres, Fontana Collins.

——— (1993), *La fortuna moral,* trad. Susana Marín, México, Instituto de Investigaciones Filosóficas-UNAM.

Wittgenstein, Ludwig (1953), *Philosophical Investigations* trad. G.E.M. Anscombe, Oxford, Basil Blackwell.

——— (1988), *Investigaciones filosóficas,* trad. Alfonso García Suárez y Ulises Moulines, Barcelona, Crítica/Instituto de Investigaciones Filosóficas-UNAM.

Wollheim, Richard (1984), *The Thread of Life,* Cambridge, Cambridge University Press.

van Straaten, Zak (comp.) (1980), *Philosophical Subjects: Essays Presented to P.F. Strawson,* Oxford, Clarendon.

ÍNDICE ANALÍTICO

Adimanto 93
adulación 194, 212-227
albedrío, libre 229, 232-235, 247
anhelo 57-58, 78
Anscombe, G.E.M. 34, 38, 41, 53, 62, 81-82, 141, 231-232, 251
Aristóteles 48, 56
Aubrey, John 153

Baker, Judith 97, 237, 238, 251
bienestar 48, 49, 50, 51, 78, 183, 206, 244
Bierce, Ambrose 218, 223
Blackburn, Simon 20, 251
Brentano, Franz 166-167, 251
Buñuel, Luis 214
Burke, Edmund 75
Butler, Joseph 75

carencias (faltas) 47-52, 60
causación 27, 36, 39, 134-135, 137-139, 143, 144-145
  razonable 72-73, 100-101, 137-150, 204-205
Cervantes, Miguel de 28-29, 249
Collingwood, R.G. 12
concepciones 70, 148-150, 175-183, 195
  del deseo 25-46, 54-58
  del deseo como poder activo 27-40, 42-46, 47, 54-55, 57, 65, 68-69, 83-84, 157
conceptos esencialmente debatibles 176-178

«dar lugar a» *véase* causación razonable
Davidson, Donald 29-36, 55-56, 93-94, 136, 138, 141, 251
deliberación protopráctica 30-32, 43-44, 59, 224
Dent, N.J.H. 75, 153, 251
deseabilidad
  atomismo acerca de la percepción de 55-57, 171
  caracterizaciones de 40-46, 52-54, 56-62, 68-74, 75-78, 80-83; *véase también* razones, uso más amplio del término
  percepciones de 52-54, 55-57, 58-62, 171-172
deseo(s) 21, 25-122, 125-126, 139-144, 146-150, 156-157
  motivados 63-77, 80, 82, 83, 90-93, 100-107, 109-119, 158
  taxonomía de los 63-122; deseos «generadores de razones» 74-76, 77, 79, 80, 81, 83, 87-89, 91, 96, 120, 157; simple querer 78-83, 89-90, 96-100, 107-109
dirección de adecuación de los estados mentales con el mundo 62-63, 65, 68-69, 83, 141-142
Dummett, Michael 11

Elster, Jon 44-45, 251
Epicteto 78
Evans, Gareth 23
evaluación 80, 87, 94, 96, 99, 100-107
exigencias morales 170-172
externalismo-internalismo 71-72, 93-107, 108, 109, 111, 120, 122, 154-158, 237-241

Falk, W.D. 93, 251
felicidad 66, 183, 201, 213, 216, 242
Foot, Philippa 17–18, 30, 65, 67, 69, 86, 87, 85–93, 102, 107, 119, 133, 180–181, 185, 199–201, 229, 236, 239, 244, 252

Gallie, W.B. 176–177, 179, 252
Gandhi, M. 220
Genet, Jean 60–61
Glaucón 93
Grice, H.P. 97, 101, 105–107, 237
gusto, juicios de 180–181, 200

Hampshire, Stuart 19, 252
Harman, Gilbert 188–196, 205–206, 252
Herbert, George 163
Hobbes, Thomas 153, 210
Holmes y Rahe, 148
honor, códigos de 167, 214–217, 221, 243
Hume, David 20, 21, 108–109, 114–117, 119, 121–122, 125–161, 168, 184, 191–195, 209, 210, 224, 247, 252–253

ideas, teoría de las 126–127, 132–133, 135–136, 137, 143, 144–145
internalismo-externalismo *véase* externalismo-internalismo

Johnson, Samuel 48
juicios internos 189–196
justicia 177, 182–183, 206, 244
Justiniano 167, 174

Kant, Immanuel 55, 173, 192, 202, 235, 237, 246, 253
Kenny, Anthony 140, 253
Kraus, Karl 13, 47

Law, William 211, 217–219, 253

MacIntyre, Alasdair 17, 153, 169, 199, 226, 235, 244, 253
Mackie, J.L. 20, 151, 155, 160, 253
Mandeville, Bernard de 209–227, 241–242, 247, 249, 253
Maquiavelo, Nicolás 47
Maupassant, Guy de 25
McDowell, John 11, 56, 66–73, 77, 91, 114–116, 120, 169, 253
McGinn, Colin 108, 253
Mill, John Stuart 73–74, 253
Monro, Hector 223, 254
moralidad
  como cuestión de hecho 125–161, 214–215, 217–220, 235–241
  escepticismo acerca de la 13–22, 120–121, 217, 220, 222–224, 227–228, 239, 247
  metafísica descriptiva de la 19–22
  naturaleza impráctica de la 214–215, 217–220, 235–241
  naturaleza práctica de la 130, 133, 151–161, 164–165
«moralidad sexual» 13, 151, 167, 185, 222
moralismo 173, 203–204, 236, 240–241, 245–247
motivación moral 121–122, 146–150, 152–154, 156–161, 188–196, 215–216, 217, 219, 220, 237–241, 247–248
muerte 147–148
Murdoch, Iris 123, 165–166, 239–240, 254

Nagel, Thomas 63–73, 74, 77, 78, 87, 90, 91, 120, 158, 159, 163–164, 254
naturaleza
  humana 185, 210–213, 216–218, 223, 225–226, 229, 230, 239–240, 247–248
  moral del mundo 223–224, 235–236, 240–241

necesidades 47–52, 60, 111, 181, 184
y carencias básicas 48, 50–52
Nietzsche, Friedrich 123, 209, 227–248, 254
Nozick, Robert 51, 244, 254

objetividad y valor *véase* subjetividad y valor
O'Shaughnessy, Brian 38, 254

pasiones apacibles 132–133, 134, 156–157
Peacock, Thomas Love 11, 13
Pears, David 37–38, 254
placer 74–76, 126, 138–139, 152–154, 156–159, 212, 220, 223, 238–239
Pope, Alexander 48
Private Eye 49
prudencia 64–66, 93, 168
Putnam, Hilary 51, 255

razones
para actuar 35, 54, 63–69, 129, 132–134, 144–145, 188–196
uso más amplio del término 165–166, 191–194, 199, 201–203, 206–207
y caracterizaciones de deseabilidad 68–71, 78–80, 102–103, 110–112, 147–150, 206–207
Reid, Thomas 23, 159, 255
relativismo moral 183–207, 247
Richards, David A.J. 169, 255
Rochefoucauld, François La 210
Russell, Bertrand 25–36, 40–46, 60, 244, 255

Samuels, Alec 85
Sasía, Martha 11
Schiffer, Stephen 74–76, 77–78, 80, 87, 255
Shakespeare, William 85
Simonton, Matthews y Creighton 148
Smith, Michael 63, 255
Snowdon, Paul 11
Sócrates 93, 193
Spinoza, Baruch de 83
Stout, Rex 23
Strawson, P.F. 19, 125, 168, 232, 233–235, 243–244, 246, 255
Stroud, Barry 135–136, 139, 144, 145, 152, 255
subjetividad y valor 50–51, 73–74, 83–84, 107–122, 170–172; *véase también* subjetivismo moral
subjetivismo moral 20, 21, 108, 116–117, 121–122, 125, 161, 183–207
Surtees, R.S. 163
sustitución, actividades de 60, 61, 63, 69, 82, 83

Taylor, Richard 42–46, 53, 58–59, 255
teorías externas y teorías internas 15–19, 234–235
Tomás de Aquino, santo 33, 40–41, 43, 45
trascendentalismo, 73, 77–80, 114–119

Urmson, J. 55, 170, 255

vida, sentido de la 42, 86, 166, 169–170, 229

Warburton, William 125, 209
Warnock, G.J. 169, 256
Wiggins, David 45–46, 47, 116, 117, 166, 175, 184, 256
Williams, Bernard 17–18, 78–79, 167, 172, 184, 191, 193, 196–199, 202, 256
Wittgenstein, Ludwig 39, 165, 256
Wollheim, Richard 58, 256

Realidades morales. Ensayo sobre psicología filosófica se terminó de imprimir y encuadernar en los talleres de Programas Educativos, S.A. de C.V., Calz. Chabacano 65, local A, col. Asturias, México, D.F., el 25 de marzo de 1998. Se tiraron 2000 ejemplares. En su composición y formación, realizadas por computadora en el Instituto de Investigaciones Filosóficas, se utilizaron el programa TEX y tipos New Baskerville 10/12, 9/11 y 8/10; cuidó la edición Laura E. Manríquez.